

FIGURE 1

GGGGCTCGCGGCCAGCGGCCAGCGCTAGTCGGCTGGTAAGGATTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTGTGAAGTTGTAACAGAAAACCTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCTCAGTTCCCTTCAGCCCTGTAATTGGACATCTGCTGCTTCATATTT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTGGGCAATGCTAAATATTGGCAGTTATGCATTG
CTACCATTATGTTGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAAA
TTAAACAAGGCTGGCCTTGTACTTGAATACTGAGTTGTTAGGACTTCTATTGGCAAACCTT
CCAGAAAAACAACCCCTTTGCTGCACATGTAAGTGGAGCTGTGCTTACCTTGGTATGGCCTCAT
TATATATGTTGTTGACCATCCTTCCACAAATGCAGCCAAACATGGCAAACAGTC
TTCTGGATCAGACTGTTGGTATCTGGTGGAGTAAGTGCACCTAGCATGCTGACTTGCTC
ATCAGTTTGACAGTGGCAATTGGACTGATTAGAACAGAAACTCCATTGGAACCCGAGG
ACAAAGGTTATGTGCTTCACATGACTACTGCAGCAGAATGGCTATGTCATTCCCTTT
GGTTTCTGACTTACATTGATTTAGAAACAGAAACTCCATTGGAACCCGAGG
TGGATTAACCCCTCATGACACTGCACCTGCCCTATTAAACATGAACGAACACGGCTACTTCCA
GAGATATTTGATGAAAGGATAAAATATTCTGTAATGATTGATTCTCAGGGATTGGGAAAGG
TTACAGAAGTTGCTTATTCTCTGAAATTCAACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAAGGCCATTGATAGATTATTCTAAAGGATATCAT
CAAGAAGACTATTAAAACACCTATGCCTATACCTTTATCTCAGAAAATAAGTCAAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIFSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQVFIRLLLVIWCGVSALSMLTCSSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFLTYIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACCGCGTGGCGGACCGTGGGGAGAGCCGAGTCCCGCTGCAGCACCTGGAGAAGGCAGACC
GTGTGAGGGGGCCTGTGGCCCCAGCGTGTGGCCTGGGGAGTGGAAAGTGGAGGCAGGAGCCTTC
CTTACACTTCGCCATGAGTTCCATCGACTCCAGCATCATGATTACCTCCAGATACTATTTTTG
GATTTGGGTGGCTTCTCATGCGCAATTGTTAAAGACTATGAGATACTGTCAGTATGTTGACAG
GTGATCTTCTCGTGACGTTGCATTTCCTGCACCAGTGTGACTCATCATCTTGAAATCTTAGG
AGTATTGAATAGCAGCTCCCGTTATTCACTGGAAAATGAACCTGTGTAATTCTGCTGATCCTGG
TTTCATGGTGCCCTTACATTGGCTATTGTTATTGTGAGCAATATCCGACTACTGCATAAACAGA
CTGCTTTTCTGTCTTATGGCTGACCTTATGTATTCTGGAAACTAGGAGATCCCTTCC
CATTCTCAGCCCCAAACATGGATCTTATCCATAGAACAGCTCATGCCGTTGGTGTGATTGGAG
TGACTCTCATGGCTCTTCTGGATTGGTGTGCACTGCCATACACTACATGTCTTACTTC
CTCAGGAATGTGACTGACACGGATTCTAGCCCTGGAACGGGACTGCTGCAAACCATGGATATGAT
CATAAAGCAAAAGAAAAGGATGGCAATGGCACGGAGAACAAATGTTCCAAGGGGAAAGTGCATAACA
AACCATCAGGTTCTGGGAATGATAAAAGTGTACCACTCAGCATCAGGAAGTGAATCTTACT
CTTATTCAACAGGAAGTGGATGCTTGGAAAGAATTAAGCAGGCAGCTTTCTGAAACAGCTGATCT
ATATGCTACCAAGGAGAGAATAGAAACTCCAAAACCTCAAGGGAAATATTTAATTCTGGTT
ACTTTTCTCTATTACTGTGTTGGAAAATTTCATGGCTACCATCAATATTGTTTGATCGAGTT
GGGAAAACGGATCTGTACAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTGATGT
GAAGTTTGGTCCCAACACATTCTCTCATCTGGGAAATCATCGTCACATCCATCAGAGGAT
TGCTGATCACTCTTACCAAGTCTTTATGCCATCTCTAGCAGTAAGTCCCAATGTCATTGCTCTG
CTATTAGCACAGATAATGGCATGTACTTGTCTCTGTGCTGCTGATCCGAATGAGTATGCCTT
AGAATACCGCACCAATCACTGAAGTCCTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGTTG
ATGTGATCTCCTGGTCAGCGCTCTCTAGCATACTCTCTGATTTGGCTACAAACAGGCACCA
GAGAAGCAAATGGCACTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTCAAAATTAA
GATATAAGAGGGGGAAAATGGAACCAGGGCTGACATTATAAACAAACAAATGCTATGGTAGC
ATTTTCACCTTCATAGCATACTCTTCCCCGTAGGTGATACTATGACCATGAGTAGCATGCCAG
AACATGAGAGGGAGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCGTGTGGATATGAGGCTGG
TGTAGAGGCGGAGAGGAGCCAAGAAACTAAAGGTGAAAATACACTGGAACCTGGGGCAAGACATGT
CTATGGTAGCTGAGGCCAACACGTTAGGATTCCGTTTAAGGTTCACATGGAAAAGGTTAGCTTGC
CCTTGAGATTGACTCATTAAATCAGAGACTGTAACAAAAAAAAAAAAAAGGGCGCCGCG
ACTCTAGAGTCGACCTGCAGAAGCTGGCCCATGGCCAACCTGTTATTGCAGCTTATAATG

FIGURE 4

MSFLIDSSIMITSQLFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVDFAFSCTMFELIIFEILGV
LNSSSRYFHWKMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFFWKLGD
FPILSPKKGILSIEQLISRVGIVGVTLMALLSGFGAVNCPTYMSYFLRNVTDTDILALERRLLQ
TMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWMKSVTTSASGSENLTLLIQQEVDALEELSRQ
LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTDPVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLAQIMGMY
FVSSVLLIRMSMPLYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCTCGTTATGAAGTGGAGCAGTGAGTGTGAGCCTAACATAGTTCC
AGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAAGTGGCCATCTGAGGT
GTTTCCCTGGCTCTGAAGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT
TCCATCTGGACCACGAGGCTCTGGTCAAGGCTTTGCGTGCAGAAGAGCTTCCATCCAGGT
GTCATGCAGAATTATGGGGATCACCCCTGTGAGCAAAAGGCAGCACGAGCTGAATTTCACAG
AAGCTAAGGAGGCTGTAGGCTGCTGGGACTAAGTTGCCGGCAAGGACCAAGTTGAAACAGCC
TTGAAAGCTAGCTTGAACATTGCACTATGGCTGGGTGGAGATGGATTGCGTGTATCTCTAG
GATTAGCCAAACCCCAAGTGGGGAAAAATGGGGTGGGTGCTGATTGAAAGGTTCCAGTGA
GCCGACAGTTGCACTTGTGACTACACTGATACTTGGACTAACTCGTGCATTCCAGAA
ATTATCACCAACAAAGATCCCATATTCAACACTCAAACACTGCAACACAAACAGAATTATTGT
CAGTGACAGTACCTACTCGGTGGCATCCCCACTCTACAATACCTGCCCTACTACTACTCCTC
CTGCTCCAGCTTCACTTCTATTCCACGGGAGAAAAAAATTGATTGTCACAGAAGTTTATG
GAAACTAGCACCATGTCTACAGAAACTGAACCATTTGTGAAATAAAGCAGCATTCAAGAATGA
AGCTGCTGGGTTGGAGGTGCCCCACGGCTCTGCTAGTGCTGCTCTTGTGCTG
CAGCTGGCTTGGATTGCTATGTCAAAGGTATGTGAAGGCCCTCCCTTTACAAACAAGAAT
CAGCAGAAGGAAATGATGAAACCAAAGTAGTAAAGGAGGAGAAGGCCATGATAGCAACCTAA
TGAGGAATCAAAGAAAATGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAACCTACCGTGC
GATGCTGGAAGCTGAAGTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTCTTT
CATGCTCTTACCTGCCAGCTGGGAAATCAAAGGGCAAAGAACCAAAGAACAGAAAGTCCA
CCCTGGTTCTAACTGGAATCAGCTCAGGACTGCCATTGACTATGGAGTGCAACCAAAGAGAAT
GCCCTCTCTTATTGTAACCTGTCGGATCTATCCTCTACCTCCAAAGCTTCCCACGGCT
TTCTAGCCTGGCTATGCTCTAAATATCCCCTGGAGAAAGGAGTTTGCAAGTGCAAGGAC
CTAAAACATCTCATCAGTATCAGTGGTAAAAGGCCCTCTGGCTGTCTGAGGCTAGGTGGTTG
AAAGCCAAGGAGTCACTGAGACCAAGGCTCTCTACTGATTCCTACTGATTCCTGAGGCTAGGTGGTTG
GCTCTGAAAGAGAAACACGTATCCCACCTGACATGTCCTCTGAGGAGATTCTCATAACTGAGACCTAATCTGTAAA
GGCAGAAAAGTTAGCCCCCTGAAAGCCATGAGGAGATTCTCATAACTGAGACCTAATCTGTAAA
GCTAAAATAAAGAAATAGAACAGGCTGAGGAGACGACAGTACACTGTCAGCAGGGACTGTAAC
ACAGACAGGGTCAAAGTGTCTCTGAAACACATTGAGTTGGAATCACTGTTAGAACACACACA
CTTACTTTCTGGCTCTACCACTGCTGATATTCTCTAGGAAATATACTTTACAAGTAACA
AAAATAAAAATCTTATAAATTCTATTCTATCTGAGTTACAGAAATGATTACTAAGGAAGATT
ACTCAGTAATTGTTAAAAGTAATAAAATTCAACAAACATTGCTGAATAGCTACTATATGTC
AAAGTGTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCTCAAAAATGCACTAGTAG
AACGCTATCTGGGAAGCTATTCTCTGAGTTGATATTCTAGCTTATCTACTTCCAAACTAAT
TTTATTCTCTGAGACTAATCTTATTCTATTCTCTAATATGCCAACATTATAACCTTAATT
TATTATTAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAGCACATTAAAAGGCC
ATTAACAAATGTATCACTAGCCCTCTTCAACAGAACAGGACTGAGAGATGCAGAAATATT
TGTGACAAAAATAAAGCATTAGAAAACCTT

FIGURE 6

MARCFSLVLLTSIWTRLLVQGSLRAEELSIVSCRIMGITLVSKKANQQLNFTAKEACRLLG
LSLAGKDQVETALKASFETCSYGVGDGFVVISRISPNSPKCGKNGVGVLIWKVPVSRQFAAYCYN
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTPPAPASTSIPR
RKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVALLFFGAAAGLGFCYVK
RYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKTDKNPEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGCGCTCCCGCACCGCGGGCCGCCACCGCGCCGCTCCGCATCTGCACCCGAGCCCGC
GGCCTCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCGAGCGAACCTCGTCCAGTCGGCG
CGGCTGCGGGCGCAGAGCGGAGATGCAGCGGCT TGGGGCCACCCCTGCTGTGCTGCTGGCG
CGGCGGTCCCCACGGCCCCCGGCCGCTCCGACGGCGACCTCGGCTCCAGTCAAGCCGGCCG
GCTCTCAGTACCCGAGGAGGAGGCCACCTCAATGAGATGTTCCGCAGGTTGAGGAACGTGAT
GGAGGACACGCAGCACAAATTGCGCAGCGGGTGAAGAGATGGAGGCAGAAGAAGCTGCTGCTA
AAGCATCATCAGAAAGTGAACCTGGAAACTTACCTCCAGCTATACAATGAGACCAACACAGAC
ACGAAGGTTGGAATAATACCATCCATGTGACCGAGAAATTACAAGATAACCAACACAGAC
TGGACAAATGGTCTTTCAGAGACAGTTATCACATCTGTGGGAGACGAAGAAGGCAGAAGGAGCC
ACGAGTGCATCATCGACGAGGACTGTGGGCCAGCATGTA
CTGCAGTGTGCTGGAGACCA
GCTGTGCTGGGACTGCTGACCAAAATGGCCACCAAGGGGAGCAATGGACCATCTGTGACA
ACCAAGGGACTGCCAGCCGGGCTGTGCTGTGCTTCCAGAGAGGCTGTGTTCCCTGTGTC
ACACCCCTGCCGTGGAGGGCAGCTTGGCATGACCCGCCAGCCGCTTCTGGACCTCATCAC
CTGGGAGCTAGAGCCTGATGGAGCCTGGACCGATGCCCTGTGCCAGTGGCTCCCTGCCAGC
CCCACAGCCACAGCCTGGTGTATGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
GAGATCCTGCTGCCAGAGAGGGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
CCAGGAGCTGGAGGACCTGGAGGAGGAGCCTGACTGAAGAGATGGCCTGGGGAGCCTGCGCTG
CCGCGCTGCACTGCTGGAGGGGAAGAGATTTAGATCTGGACCAGGCTGTGGTAGATGTGCAA
TAGAAATAGCTAATTATTCAGGTGTGCTTAGGCGTGGCTGACCAGGCTTCTCAGGAGGAGGAGG
CATCTCTTCCCAGTAAGTTCCCTCTGCTTGACAGCATGAGGTGTGCAATTGTTCA
CCCCCAGGCTGTTCCAGGCTTCACAGTGTGCTGGGAGAGTCAGGCAGGGTTAAACTGCA
GGAGCAGTTGCCACCCCTGTCCAGATTATTGGCTGCTTGCCTTACCAAGTTGGCAGACAGCCG
TTTGTCTACATGGCTTGTATAATTGTTGAGGGGAGGAGATGGAAACAATGTGGAGTCTCCCTC
TGATTGGTTGGGAAATGTGGAGAAGAGTGCCTGTTGCAAACATCAACCTGGAAAATG
CAACAAATGAATTTCACGCA
GTTCCAGGCTTCCATGGCATAGCTGTGCTGGGAGAGTCAGGCAGGGTTAAACTGCA
AGATGAAATGTTCTGTTACCCCTGCATTACATGTGTTATTCA
TCCAGCAGTGTGCTCAGCTCC
TACCTCTGTGCCAGGGCAGCATTTCATATCCAAGATCAATTCCCTCTCTCAGCACAGCCTGGG
AGGGGGCATTGTTCTCTCGCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTCATCTGGTGTGACTCTAAGCTCAGTGTCT
CTCCACTACCCACACCAGCCTGGTGCACCCAAAAGTGCTCCAAAAGGAAGGAGAATGGGAT
TTTCTTGAGGCATGCACATCTGGATTAGGTCAA
ACTATTCTCACATCCCTCTAAAGTAAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGAGCCGTCTTAATGAAGACAATGAT
ATTGACACTGTCCCTTTGGCAGTTGCATTAGTA
ACTTTGAAAGTATATGACTGAGCGTAGCA
TACAGGTTAACCTGCAGAAACAGTACTTAGGTATTGAGGGCAGGATTATAATGAAATTG
AAAATCACTTAGCAGCAACTGAAGACAATTATCAACCA
CAGTGGAGAAAATCAAACCGAGCAGGGC
TGTGAAACATGGTGTAAATATGCGACTGCGAACACTGA
ACTCTACGCCACTCCACAAATGATG
TTTCAGGTGT
CATGGACTGTTGCCACCATGTATT
CATCCAGAGTCTTAAAGTTAAAGTGTATA
AACATAAGTTGCATTAGAA
ATCAAGCATAAAACTCAACTGCA
AAAAAAAAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTOHKL
RSAVEEMEAEEAAKASSEVNLNLPPSYHNETNTDTKGNNTIHVREIHKITNNQTGQMVFSE
TVITSGVDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRQMLCTRSECCGDQLCVWGHC
TKMATRGSNGTICDNQRDCQPGLCCAFQRGLLFPVCTPLPVEGELCHDPASRLLDLITWELEPDG
ALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQELEDLE
RSLTEEMALGEPAAAAALLGGEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGCGGACCGTGGGGCTGTGAGAAAGTCCAATAAACATCATGCAACCCAC
GGCCCACCTTGTGAACTCCTCGCCAGGGCTGATGTGCGTCTCCAGGGCTACTCATCCAAG
GCCTAATCCAACGTTCTGTCTCAATCTGCAAATCTATGGGTCCCTGGGCTTCTGGACCCCTT
AACTGGGTACTGGCCCTGGCCAATGCGTCCTCGCTGGAGCCTTGCCTCCTACTGGCCTT
CCACAAGCCCCAGGACATCCCTACCTCCCTTAATCTCGCCTTCATCCGACACTCCGTTACC
ACACTGGGTCAATTGGCATTGGAGCCCTCATCCTGACCCTGTGCAGATAGCCCAGGTACATCTG
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCGCTGTAGCCCGCTGCATCATGTGCTGTTT
CAAGTGCTGCCTCTGGTGTCTGGAAAAATTATCAAGTTCTAAACCGCAATGCATACTCATGA
TCGCCATCTACGGGAAGAATTCTGTCTCAGCCAAAATGCGTTCATGCTACTCATGCGAAAC
ATTGTCAAGGGTGGTCGTCTGGACAAAGTCACAGACCTGCTGCTGTTCTGGAGCTGCTGGT
GGTCGGAGGCGTGGGGCTCTGCTTCTGGACCCATCCGGGCTGGTAAAG
ACTTTAAGAGCCCCACCTCAACTATTACTGGCTGCCATCATGACCTCCATCCTGGGGCTAT
GTCATGCCAGCGGCTCTCAGCTTTGGCATGTGTGGACAGCCTCTGCTTCTGCTTCT
GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGCCCTACTACATGTCCAAGAGCCTCTAA
AGATTCTGGCAAGAAGAACGAGGCGCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGG
CCCTGATCCAGGACTGCACCCCACCCCCACCGTCCAGGCATCCAACCTCACTTCGCTTACAGGT
CTCCATTGTGGTAAAAAAAGTTTAGGCCAGGCCGTGGCTACGCCTGTAATCCAACACT
TTGAGAGGCTGAGGCCGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGCCAACATGGTG
AAACCTCCGTCTTATTAAAAATACAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCACTCCA
GCTACTCGGGAGGCTGAGGCAGGAGAACGCTGAACCCGGAGGCAGAGGTTGCAGTGAGCCGA
GATCGGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAA
AAGATTTATTAAAGATTTGTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVGLFWTL
NWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN
IVRVVVLDKVTDLFFFKGKLLVVGGVGVLSSFFFSGRIPGLGKDFKSPHLNYYWLPIMTSILGAY
VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYYMSKSLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

CCCCCGCGCCCGCGCCGGCGCCCAGAGCCGGGAGCCACCGCCATGGGGGCCTGCCCTGGAGCCTGC
TCCCTGCTCAGCTGCGCTCCTGCCCTGCGCTCTGCCCTGCATCCTGTGCAGCTGCTGCCCGC
CAGCCGCAACTCCACCGTGAGCCGCTCATCTTCACGTTCTTCCTCTTCCGGGGTGTGGTGTCCA
TCATTATGCTGAGCCGGCGTGGAGAGTCAGCTACAAGCTGCCCTGGGTGTGAGGAGGGGCC
GGGATCCCCACCGCCTGCAGGGCACATCGACTGTGGCTCCCTGCTGGCTACCGCGCTGTCTACCG
CATGTGCTCGCCACGGCGGCCCTCTTCTTACCCCTGCTCATGCTTGCGTGGAGCAGCA
GCCGGGACCCCGGGCTGCCATCCAGAATGGGTTTGGTTCTTAAGTCCGTACCTGGTGGGCC
ACCGTGGGTGCCTTACATCCCTGACGGCTCCCTCACCAACATCTGGTCTACTTCGGCGTCGTGGG
CTCCTTCCCTTCATCCTCATCCAGCTGGTGTGCTCATCGACTTGCACACTGGAAACCAGCGGT
GGCTGGGCAAGGCCGAGGAGTGGGATTCCCGCTGGTACCGCAGGCCCTTCTTCACTCTCCTC
TTCTACTTGCTGTCGATCGGGCGTGGCGTGTGTTCATGTAACACTGAGGCCAGCGGCTGCCA
CGAGGGCAAGGTCTTCATCAGCCTCACCTCACCTCTGTGTCTGCGTGTCCATGCTGCTGTCTGC
CCAAGGTCCAGGACGCCAGCCCCAACCTGGGCTGCTGCAGGCCCTGGTCATCACCTCACCCATG
TTTGTCACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTGCCAACCCAGCT
GGGCAACGAGACAGTTGGCAGGGCCGAGGGCTATGAGACCCAGTGGTGGGATGCCCGAGCATTG
TGGGCTCATCATCTTCTCTGTGCAACCTCTTCATCAGTCTGCGCTCCCTCAGACCACGGCAGGTG
AACAGCCTGATGCAACGGAGGAGTGCACCTATGCTAGACGCCACACAGCAGCAGCAGCAGG
GGCAGCCTGAGGGCCGGCCTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCCCTTCC
ACTTCTGCCCTGGTGTGGCTACTGCACGTCACTGATGACGCTACCAACTGGTACAAGCCGGTGAG
ACCCGGAAGATGATCAGCACGTGGACCCGCGTGTGGGTGAAGATCTGTGCCAGCTGGGAGGGCTGCT
CCTCTACCTGTGGACCCCTGGTAGCCCCACTCCTCTGCCAACCGCGACTTCAGCTGAGGCAGCCTCA
CAGCCTGCCATCTGGTGCCCTCTGCCACCTGGTGCCCTCTGGCTCGGTACAGCCAACCTGCCCTC
CCCACACCAATCAGCCAGGCTGAGCCCCACCCCTGCCAGCTCCAGGACCTGCCCTGAGCCGGC
CTTCTAGTCGTAGTGCCTTCAGGTCCGAGGAGCATCAGGCTCCTGCGCAGAGCCCCATCCCCCGCCAC
ACCCACACGGTGGAGGCTGCCCTCTCCCTCCCTCCCTGTTGCCCATACTCAGCATCTGGATGAA
AGGGCTCCCTGTCCCTCAGGCTCACGGGAGCGGGCTGCTGGAGAGAGCAGGGAACTCCACACAG
TGGGGCATCCGGCACTGAAGCCCTGGTGTCTGGTCACGTCCCCAGGGACCCCTGCCCTTCC
GACTTCGTGCCCTACTGAGTCTAAGACTTTCTAATAACAAGCCAGTGCCTGTAAAAAAA

FIGURE 12

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIIMLSPGVESQL
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFTLLMLCVSSSRDPRAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSLFILQLVLLIDFAHAWNQRWLGKAE
ECDSRAWYAGLFFFLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFCVCVSIAAVLPKV
QDAQPNGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVAGPEGYETQWWDAPS
VGLIIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQVAACEGRAFDNEQDGVTYSY
SFFHFCLVLASLHVMMTLTNWYKPGTRKMIWTAVWVKICASWAGLLYLWTLVAPLLLRND
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGCCAGGAACCACCCGTTAAGGTGTCTCTCTTAGGGATGGTGA
GGTTGGAAAAAGACTCCTGTAACCCCTCCTCAGGATGAACCACCTGCCAGAAGACATGGAGAACG
CTCTCACCGGGAGCCAGAGCTCCATGCTCTGCGCAATATCCATTCCATCAACCCCACACAA
CTCATGCCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
TTTCTGTTGTTGTCACCTTGACCTTATTGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAACGGAGGTGATGCAGTATGACTACTATTCTTCATAT
TTTGATATATTCTCTGGCAGTTTCGATTTAAAGTGTAAACTTGCAATATGCTGTGCGAG
ACTGCGCCATTGGTGGCAATAGCGTTGACAACGGCAGTGACCAGTGCTTTACTAGCAAAG
TGATCCTTCGAAGCTTTCTCTCAAGGGGCTTTGGCTATGTGCTGCCATCATTTCATTTCATC
CTTGCCTGGATTGAGACGTGGTCCTGGATTCAAAGTGTACCTCAAGAACAGAAGAAGAAAA
CAGACTCCTGATAGTCAGGATGCTTCAGAGAGGGCAGCACTTACCTGGTGGCTTGTGATG
GTCAGTTTATTCCCCCTCTGAATCCGAAGCAGGATCTGAAGAACGCTGAAGAAAAACAGGACAGT
GAGAAACCACTTAGAACATGAGTACTACTTTGTTAAATGTGAAAAACCCCTCACAGAAAGTC
ATCGAGGCAAAAGAGGCAGGCAGTGGAGTCTCCCTGTCGACAGTAAAGTTGAAATGGTGACGTC
CACTGCTGGTTTATTGAACAGCTAATAAAGATTATTGTAATACCTCACAAACGTTGTAC
CATATCCATGCACATTAGTGCCTGCCTGGCTGGTAAGGTAATGTCATGATTCACTCTCT
TCAGTGAGACTGAGCCTGATGTTAACAAATAGGTGAAGAACAGTCTGTGCTGTATTCTAAC
AAAAGACTTAATATATTGAAGTAACACTTTTTAGTAAGCAAGATACTTTTATTCAATTCA
AGAATGGAATTTTTGTTCATGTCAGATTATTTGTATTCTTTAACACTCTACATT
TCCCTGTTTTAACTCATGCACATGTGCTTTGACAGTTAAAAAGTGTAAATAAAACTG
ACATGTCAATGTGGCTAGTTTATTCTGTTGCATTATGTTGATGGCTGAAGTGTGGA
CTTGCAAAAGGGAAAGAAAGGAATTGCGAACATGTAAGAACAGTCACTGGACATTGTATT
TTATCATGAAATCATGTTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTGAAATGC
ACAAAATGACTAAACCATTCAATCATGTTCTTGCAGCCAATTCAATTAAAATGAA
CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF
VTLLWIIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQGAFGYVLPIISFILAWIETWFLDFKVLPOEAEEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCCCTGGGGCCCACCCGCCAGGAAAGACTGAGG
CCGGGGCTGCCCCGGCGCTCCCTGGCCGCCCTCCGGGACAGAACATGTGCTCCAG
GGTCCCCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGCTGGGGTGCAGGGCTGCCAT
CCGGCTGCCAGTGAGCCAGCACAGACAGTCTCTGCACTGCCGCCAGGGACACGGTGC
CGAGACGTGCCACCGACACGGTGGGGCTGTACGTCTTGAGAACGGCATACCATGCTGACGC
AGGACAGCTTGCCGGCTGGGGCCTGCAAGCTCTGACCTGTACAGAACAGATGCCAGCC
TGCCAGCGGGCTTCCAGCCACTGCCAACCTCAGCAACCTGGACCTGACGCCAACAGCTG
CATGAAATCACCAATGAGACCTCCGTGGCCTGCGGCCCTGAGCGCTCTACCTGGCAAGAA
CCGCATCCGCCACATCCAGCTGGTGCCTCGACACGCTGACGCCCTGGAGCTCAAGCTGC
AGGACAAACGAGCTGCCGGCACTGCCGCCGTGCGCCTGCCGCCCTGCTGCTGGACCTCAGC
CACAAACAGCTCTGGCCCTGGAGCCGCATCTGGACACTGCCAACAGTGGAGGCCTGCC
GGCTGGCTGGGGCTGCAAGCAGCTGGACGAGGGCTTCAAGCCGTTGCGCAACCTCCACGACC
TGGATGTGTCGACAACCAGCTGGAGCGAGTGCACCTGTGATCCGAGGCTCCGGGGCTGACG
CGCCTGCCGCTGGCCGAACACCCGCATTGCCAGCTGCCGCCAGGACCTGGCCGGCTGGC
TGCCCTGCAGGAGCTGGATGTGAGCAACCTAAGCCTGCAAGGCCCTGCCCTGGCAGCCTCTGG
TCTTCCCCCGCCTGCCGTGCTGGAGCTGCCAACCCCTTCACACTGCCGTGCCCCCTGAGC
TGGTTGGCCCCGGTGCAGAGGCCACGTACACTGCCAGCCCTGAGGAGACGCCGCTGCCA
CTTCCGCCCAAGAACGCTGCCGGCTGCTCCGGACTTGACTACGCCGACTTTGGCTGCCAG
CCACCACCAACAGCCACAGTGCCAACACAGGCCCTGGTGGGGAGGCCACAGCCTTGCT
TCTAGCTGGCTCTACCTGGCTAGCCCCACAGGCCGCAACTGAGGCCAGGCCCTGCC
CACTGCCCAACCGACTGTAGGGCTGTCCCCAGGCCAGGACTGCCACCGTCCACCTGCC
ATGGGGCACATGCCACCTGGGACACGCCACCTGGCTGTTGTGCCCGAAGGCTTCACG
GGCTGTACTGTGAGGCCAGATGGGGCAGGGACACGCCAGGCCCTACACCAGTCACGCC
GCCACACGGCTCCCTGCCCTGGCAGGCCCTGGCAGGCCACCTGCCCTGCCGCTGGGGCTGC
AGCGCTACCTCCAGGGAGCTCGTGCAGCTCAGGCCCTCCGCTCACCTATCGCAACCTATCG
GCCCTGTATAAGGGCTGGTGCAGCTGCCACTGCCCTGCCCTGCCCTGCC
GCTGCCGCCAACGCCACTTACTGCCCTGTGCTGCTGCTGCTGAGTACACGGTACCCA
GCGAGGAGGCTGCCGGGAGGGCCATACACCCCCCAGCCGCTCACTCCAAACGCCAGTCACC
CAGGCCGCCAGGCCAACCTGCCCTGCCCTATGCCGCCGCCCTGGCCGGTGCCTGCC
GCTGGCTGCCGGGAGGGCTACTGTGTCGGCGGGGCCATGGCAGCAGGCCGCTCAGG
ACAAAGGGCAGGTGGGGCAGGGCTGGGGCCCTGGAACTGGAGGGAGTGAAGGTCCCCCTGG
CCAGGGCCGAAGGCCAACAGAGGGCGGTGGAGAGGCCCTGCCAGGGCTTGAGTGTGAGGTGCC
ACTCATGGCTTCCCAGGGCTGGCCTCCAGTCACCCCTCCACGCCAACGCCCTACATCTAAGCCA
GAGAGAGACAGGGCAGCTGGGGGGCTCTCAGGCCAGTGAAGATGCCAGGCCCTCC
ACACCACTGTAAGTCTCAGTCCAACCTGGGATGTGTCAGACAGGGCTGTGACCAAGCT
GGGCCCTGTTCCCTGGAACCTCGGTCTCCTCATCTGTGAGATGCTGTGGGCCAGCTGACGCC
CTAACGGCTCCAGAACCGAGTGCCTATGAGGACAGTGTCCGCCCTGCCCAACGTGCA
CCTGGGCAGGGGGCCCTGCCATGTGCTGGTAACGCACTGCCCTGGGCTCTGCC
TCCAGGCCAACCTGGGGCCAGTGAAGGAAGCTCCCCGAAAGAGCAGAGGGAGAGGCC
GGCTGTGTGACTCTAGCTTGGCCCCAGGAAGCGAAGGAACAAAGAAACTGGAAAGGAGATGC
TTTAGGAACATTTTGTAAATATATATATTATAAGAGATCCTTCCCATTATTCT
GGGAAGATGTTTCAAACCTCAGAGCAAGGACTTTGGTTGTAAAGACAAACGATGATGAA
GGCCTTTGTAAGAAAAAATAAAGATGAAGTGTGAA

FIGURE 16

MCSRVPLLLPLLLLALGPGVQGCPGQCQCSQPQTVFCTARQGTTVPRDVPPDTVGLYVFENGIT
MLDAGSFAGLPGLQQLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRLERLY
LGKNRIRHIQPGAFDTLDRLLEKLQDNELRALPPLRPLRLLLDLISHNSLLALEPGILDtanve
ALRLAGLGLQQLDEGLFSRLRNLHLDLVDSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDL
AGLAALQELDVSNLSQLALPGDLSGLFPRIRLLAAARNPFNCVCPLSWFGPWVRESHVTLASPEE
TRCHFPPKNAGRLLLELDYADFGCPATTTATVPTTRPVVREPTALSSSLAPTWLSPTAPATEAP
SPPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGTRHHLACLCPEGFTGLYCESQMGQGTRPSPTP
VTPRPPRSLTGIEPVSPSTSLRVGLQRYLQGSSVQLRSRLTYRNLSGPDKRLVTLRLPASIAEY
TVTQLRPNATYSVCVMPLGPGRVPEGEEACGEAHTPPAVHSNHPVTQAREGNILPLLIAPALAAV
LLAALAAVGAAYCVRRGRAMAAAQDKGQVPGAGPLELEGVKVPLEPGPKATEGGGEALPSGSE
CEVPLMGFPGPGLQSPHLAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

GCAGCGCGAGGCAGCGGTGGCTGAGTCGTGGCAGAGGCAGAGCTCATGCG
GGTCCGGATAGGGCTGACGCTGCTGTGCGGTGCTGAGCTGGCCTCGCGTCTCGG
ATGAAGAAGGCAGCCAGGATGAATCCTAGATTCAAGACTACTTGACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTCTTGATTCAAAGAAATCTGA
ATTAGAATCCTTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGAAAGTGTACAG
AAGATATCAGCTTCTAGAGTCTCAAATCCAGAAAACAAGGACTATGAAGAGCCAAGAAAGTA
CGGAAACCAGCTTGACCGCATTGAAGGCACAGCACATGGGAGCCCTGCCACTTCCCTTTCT
TTTCCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAAGATGGCAGACTGTGGTGTG
CTACAACTATGACTACAAAGCAGATGAAAAGTGGGCTTTGTGAAACTGAAGAAGAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAACGTGAATGAAAATCTTAATGGAAG
CAATAAGAAAAGCCAAAAAGAGAAGCATATCGGTATCTCAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGTCAATGCTCTTATTGGTGTAAATTGCCACAGAATATCCAG
GCAGCGAGAGAGATGTTGAGAAGCTGACTGAGGAAGGCTCTCCAAGGGACAGACTGCTCTGG
CTTCTGTATGCCCTGGACTTGGTGTAAATTCAAGTCAGGCAAAGGCTTGTATATTACAT
TTGGAGCTCTGGGGCAATCTAATGCCACATGGTTGGTAAGTAGACTTTAGTGGAAAGGCT
AATAATATTAACATCAGAAGAATTGTTGTTATAGCGGCCACAACCTTTTCAGCTTCTGATGATC
CAGATTGCTGTATTAAGACCAAAATTCAAGTTGAACTTCTTCAAATTCTGTTAATGGATAT
AACACATGGAATCTACATGAAATGAAAGTTGGTGGAGTCCACAATTCTTAAATGATTAG
TTGGCTGATTGCCCTAAAAAGAGAGATCTGATAATGGCTCTTTAAATTCTGAGTTG
GAATTGTCAGAATCATTTTACATTAGATTATCATAATTAAAAATTCTTCAATTGTTGAGTTCA
AAATTGTAATGGCTATAGAAAAACACATGAAATTATAACAATTGGCAACAAATGC
CCTAAGAATTGTTAAATTCAATGGAGTTATTGTCAGAATGACTCCAGAGAGCTACTTCTG
TTTTTACTTTCATGATTGGCTGTCTCCATTATTCTGGTCAATTGCTAGTGACACTGT
GCCTGCTTCCAGTAGTCTCATTTCCATTGGCTAATTGTTACTTTCTTGCTAATTGG
AAGATTAACATTTAATAAAATTATGCTAAGATAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 18

MRVRIGLTLLLCAVLLSLASASSDEEGSQDES LDSKTTLSDESVKDHTAGR VVAGQI FLDSEES EL
ESSIQEEEDSLKSQEGESVTEDISFLESPNPNENKDYE EPPKKVRKPALT AIEGTAHGE PCHFPFLFLDK
EYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEAAKRRQMQEAEEMMYQTGMKILNGSNKKSQKR
EAYRYLQKAASMNHTKALERVSYALLFGDYL PONIQAAREMFEKLTEEGSPKGQTALGFLYASGLGVN
SSQAKALVYYTGFALGGNLIAHMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCAAGTTTAAGCCCATTCTGCAGTGGAAATTCACTGAACTAGCAAGAGGACACCATCTTCTT
GTATTATACAAGAAAGGAGTGTACCTATCACACACAGGGGGAAAAATGCTCTTGGGTGCTAGG
CCTCCTAATCCTCTGGTTCTGTGGACTCGTAAAGGAAACTAAAGATTGAAGACATCACTG
ATAAGTACATTTTATCACTGGATGTGACTCGGCTTGGAAACTTGGCAGCCAGAACTTTGAT
AAAAAGGGATTCATGTAATCGCTGCCTGCTGACTGAATCAGGATCAACAGCTTAAAGGCAGA
AACCTCAGAGAGACTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAGAGGACTG
CCCAGTGGGTGAAGAACCAAGTGGGGAGAAAAGGTCTCTGGGTCTGATCAATAATGCTGGTGT
CCCGGCGTGTGGCTCCCAC TGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA
CCTGTTGGACTCATCAGTGTGACACTAAATATGCTCCTTGGTCAAGAAAAGCTCAAGGGAGAG
TTATTAAATGTCTCCAGTGTGGAGGTGCGCTTGCATCGTGGAGGGCTATACTCCATCCAAA
TATGCAGTGGAAAGGTTCAATGACAGCTTAAGACGGGACATGAAAGCTTTGGTGTGCACGTCTC
ATGCATTGAACCAGGATTGTTCAAAACAAACTGGCAGATCCAGTAAAGGTAAATTGAAAAAAAC
TCGCCATTGGGAGCAGCTGTCTCCAGACATCAAACAATATGGAGAAGGTTACATTGAAAAAA
AGTCTAGACAAACTGAAAGGCAATAATCCTATGTGAACATGGACCTCTCCGGTGGTAGAGTG
CATGGACCACGCTCTAACAGTCTTCCCTAACACTCATTATGCCGTGGAAAAGATGCCAAA
TTTCTGGATACTCTGTCTCACATGCCAGCAGCTTGCAAGACTTTTATTGTTGAAACAGAAA
GCAGAGCTGGCTAATCCAAGGCAGTGACTCGCTAACACAAATGTCTCCTCCAGGCTATGA
AATTGGCCATTCAAGAACACATCTCCTTCAACCCATTCTTATCTGCTCCAACCTGGACT
CATTTAGATCGTGTCTATTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGGTATCCCAGGGT
CCCTGCTCAAGTTTCTTGAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCCGTGCC
GTATTTAGGCTTGCCTGCTTGGTGTGATGTAAGGAAATTGAAAGACTTGCCATTCAAAATGA
TCTTTACCGTGGCCTGCCCATGCTTATGGTCCCCAGCATTACAGTAACTTGTGAATGTTAAGT
ATCATCTCTTATCTAAATATTAAAGATAAGTCACCCAAAAAA
AAAAAAAAAAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGKLKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVIAPTDWLTLEDY
REPIEVNLFGGLISVTLNMLPLVKKAQGRVINSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
AFGVHVSCIEPGLFKTNLADPVKVIKEKKLAIWEQLSPDIKQQYGEGYIEKSLDKLKGKNSYVNMD
LSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQKAE LANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGCCGGCGGTAGCATGGAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTGTGCTCGGCG
CACTCGCTTCCAGCACCTCAACACGGACTCGGACACCGGAAGGTTCTTCTGGGAAGTAAAA
GGTGAAGCCAAGAACAGCATTACTGATTCCAAATGGATGATGTTGAAGTTGTTATAACAATTGA
CATTCAAGAAATATATTCCATGCTATCAGCTTTAGCTTTATAATTCTTCAGGCGAAGTAAATG
AGCAAGCACTGAAGAAAATATTCAATGTCAAAAGAATGTGGTAGGTTGGTACAAATTCCGT
CGTCATTCAAGATCAGATCATGACGTTAGAGAGAGGCTGCTCACAAAACCTGCAGGAGCATT
TTCAAAACCAAGACCTTGTCTGCTATTAAACCCAAGTATAATAACAGAAAGCTGCTACTC
ATCGACTGGAACATTCTTATATAAAACCTCAAAAGGACTTTTACAGGGTACCTTAGTGGTT
GCCAATCTGGGCATGTCTGAACAACTGGGTATAAAACTGTATCAGGTTCTGTATGTCCACTGG
TTTAGCCGAGCAGTACAAACACACAGCTCAAATTTTGAAAGAAGATGGATCCTTAAAGGAGG
TACATAAGATAAAATGAAATGTATGCTTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAGTG
GAAGACAGTGAACAAAGCAGTAGATAACTAGTAAAGGATGTAACAGATTAAACGAGAAATTGA
GAAAAGGAGAGGAGCACAGATTCAAGGAGCAAGAGAGAAGAACATCCAAAAGACCCCTCAGGAGA
ACATTTCCTTGTCAAGCATTACGGACCTTTTCAAATTCTGAATTCTTCATTCACTGTGTT
ATGTCTTAAAAAATAGACATGTTCTAAAGTAGCTGTAACCTACAACCACATCTGATGTAGT
AGACAATCTGACCTAATGGAGAACACACTGACATTCTGAAGCTAGTCCAGCTAGTACACCAC
AAATCATTAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTGGTTGTTA
GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTACCAAGATAAACATCCAAAAT
GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTGGTGAATATTCACGGCTC
CTACATTTTGATCCTTTAACCTTACAAGGAGATTTTTATTGCTGATGGTAAAGCCAAAC
ATTCTATTGTTTACTATGTTGAGCTACTGCAGTAAGTCATTGTTTACTATGTTCA
TGTTGAGCTAATACACAGATAACTCTTAGTGCATTACTTCACAAAGTACTTTCAAACATCA
GATGCTTTATTCCAAACCTTTTACCTTCACTAAGTTGAGGGAAAGGCTTACACAG
ACACATTCTTAAAGGAAAGTGGAGGACAGGAGGACAGTGGCTCACACCTGTAATCCCAGCACT
TAGGGAAGACAAGTCAGGAGGATTGATTGAAGGCTAGGAGTTAGAGACCAGCCTGGCAACGTATT
GAGACCATGTCTATTAAAAAATGAAAAGCAAGAATAGCCTTATTTCAAAATATGGAAA
GAAATTATGAAAATTCTGAGTCATTAAATCTCCTTAAGTGTAAACTTTTAAAGTGAAGTA
CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAATTGCAAAACATCATCT
AAAATTAAAAAAAAAAAAAAAAAAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTDSDEGFLLGEVKGEAKNSITDSQMDDVEVVYTIDIQKYI
PCYQLFSFYNSSGEVNEQALKKILSNVKNNVGWYKFRRHSDQIMTRERLLHKNLQEHFSNQDL
VFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSRAV
QTHSSKFFEDGSLKEVHKINEMYASLQEELKSICKVEDSEQAVDKLVKDVLKREIEKRRGA
QIQAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCNYNHLDVVVDNLT
MVEHTDipeaspastPQIIKHKALDLDDRWFQFKRSRLLDTQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGGAGGGCAGAGTCAGCCGAGCGAGTCAGCCGGACAGCGCAGGGCAGCCAA
GCAGCCGCAGCGAACGCCGCGCCGCCACACCCCTGCGGCTCCCGCGCGCTGCCACCCCTCCCTCCCT
GCGTCCCCGCTCGCCGGCAGTCAGCTTGCGGGGTTGCGTGCCTGCCCGCGAACCCGAGGTACCCAGCCGC
GCTTCCCTGGGCCGCGCCGCCCTCCACGCCCTCCCTGCCCTGCCCGCGCTGGCACCGGGACCGTTGCC
CGCAGGCCAGCTACTTTGCCCGCTCTCCCGCTGCTGCCCTTCCACCAACTCCA
TCCAGCTTCACTCGTAGTCCCAGCTCGCCAGCCCTGCCCGCTGCCGTAGGCCGCTCCCGTCCGGT
GGTGGGAACGCGTCCGCCCGGCCGACCAATGGCAGGTTCGGCTTGCCCGCTCTG
AGCGCCGCGCTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGGAAGTCGACGTCTTACGTG
AACAGATGCCATTCCACAGAGATCAACGGTATCATTGAAGATCTGCCAGGGTCTACCTGCT
CAAGAGATGGAGGAAGTACAGCTGCAAAGTAAAGATGATTCAAAGTGTGGT
AGCGAACAGTGAATCATTG
CAAGCTGCTTGCTTACGTTACAAGAAGTTGATGAATTCTCAAAGAACTACTTG
AAATGAGATGTTGTAAGACATATGCCATTATACATGCAAATCTGAGCTATTAAAGATCTCTCG
AAACGTTACTACGTGGTGGAAATGTAACCTGGAAGAAATGCTAAATGACTCTGG
CTCGCCCTGGAGCGGATG
TTCCGCCTGGTGAACCTCCAGTACCACTTACAGATGAGTATCTGGAATG
GTGAGCAAGTATACGGAGCAGCTGAAG
CCCTTCGGAGATGTCCTCGCAAATTGAAAGCTCCAGGTTACTCGT
CTTGTAGCAGCCCTACTTCGCTCAAGGC
TTAGCGGTTGCGGGAGATGTCGTGAGCAAGGTCTCGTGGTAAACCC
CACAGCCCAAGTGTACCCATGCCCTGTTGAAG
ATGATCTACTGCTCCCACTGCCGGGTCTCGTGA
ACTGTGAAGCCATGTTACA
ACTACTGCTCAAACATCATGAGAGGC
TGTTTGGCCAACCAAGGGATCTGATTTGA
ATGAA
CAATTTCATAGATGCTATGCTGATGGTGGCAGAGAGGGCTA
GAGGGTCTTCAACATTGAATCGT
CATGGATCCCATCGATGTAAGATTCTGATGCTATTATGAACATGCAGGAT
AAATAGTGTCAAGTGTCTCAGAAGGTTTCCAGGGATG
GGACCCCAAGCCCTCCAGCTGG
GACGAATTCTCGT
TCCATCTGAAAGTGCCTTCAGTGC
CTCGTCA
GAGCAGGAACTGGTACTGATGTC
CAAGGAGAACTGAAACAGGCCAAGAAATTCTGG
CTCCCTCCAGCTGG
GACGAATTCTCGT
AGTTGGACCGACTGGTACTGATGTC
CAAGGAGAACTGAAACAGGCCAAGAAATTCTGG
CTCCCTCCAGCTGG
GACGAATTCTCGT
GTTTGCAACGATGAGAGGATGGCTGC
AGGAAACGCCAATGAGGATGACTGTT
GGAAAGGGCAAAAGCAGGTAC
CTGTTGCA
GAGCAGGAAATGGATTAGCA
ACCCAGGGCAACAACCCAGAGGT
CCAGGTTGACACCAGCAA
ACAGAC
ATACTGATCCTCGTCAA
ATCGCTCTCGAGT
GATGACCA
GAGAATGCA
GATGCA
GACTTCTTGA
TATCAGT
GATGAAAGT
AGTGGAGAAGGAAGT
GGAGTGG
CTGTGAGT
ATCAGCAGT
GCCCTTCAGAG
TTTGACT
ACAATGCC
ACTGAC
CATGCT
GGCA
CAGGCT
ACCTCT
ACTGCT
CGATCT
GATCT
GTTCT
GGTTAT
GCA
GAGAGAG
ATGGAGA
TAATT
CTCAA
ACTCTGAG
AAAAAGTGT
CATCA
AAAAGT
TAAAGG
CACCAG
TTATCA
CTTTCT
ACCATC
CTAGT
GACTT
GGCTT
AAATGAA
TGG
GACA
AAATG
TACAG
TTTACT
ATG
GGCC
ACTGG
TTAAG
AAGT
GCT
GACTT
GTTCT
CATT
CAG
TTGG
AGG
AAAAGGG
ACTGT
GCATT
GAGTT
GGT
CCTG
CTCCCC
AAACCA
ACCATG
TTAAAC
GTT
GGCT
AACAG
GT
TAGGT
ACAGA
CT
ATAGT
TAGT
GTG
ATT
GTG
ATT
TCA
ACT
CT
TATT
TTG
TTG
TAT
GTT
TTT
CT
CAT
TT
CGTT
GTGG
TT
GG
TTT
CCA
ACT
GTG
AT
CTCG
CCT
GTT
CT
ACA
GCA
AAC
AGGG
TCC
CTT
GG
CAC
GTA
AAC
ATG
TAC
GT
AC
TATT
TCT
ATT
TAT
GTT
TAT
TAA
AG
AAAAG
GCC
AAAAG
GC

FIGURE 24

MARFGLPALLCTLAVLSAALLAELKSKCSEVRRLYVSKGFKNDAPLHEINGDHLKICPQGST
CCSQEMEEKYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSNDMFVKTYGH
LYMQNSELFKDLFVELKRYYVGVNLEEMLNDFWARLLERMFRLVNSQYHFTDEYLECVSKYTE
QLKPFGDVPRKLQLQVTRAFVAARTFAQGLAVAGDVVKVSVNPTAQCTHALLKMIYCSHCRLG
VTVKPCNYCSNIMRGCLANQGDLDFEWNNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIMN
MQDNSVQVSQKVFQGCGPPKPLPAGRISRSISESAFSARFRPHHPEERPTTAAGTSLDRLVTDVK
EKLKQAKKFWSLPSNVNCNDERMAAGNGNEDDCWNGKGKSRYLFAVTGNGLANQGNNPEVQVDT
KPDILILRQIMALRVMTSKMKNAYNGNDVFFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
CTGAGTCATCCCCAGGGATCAGGGCCTCCAGCAGGGAACCTCATTATATTCTTCAAGCAACT
TACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTCTCCCTCCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTAGCAGCCTGAATCCAGGGTCGCCAGAGGCCACAGGGACCAGGC
CAGGCTCTAGGAGATGGCTCCAGGAAGGGCCAAGAATGTGAGTGAAAGATTGGTCTGAG
AGCCCCGAGAAGAAAATTCACTGACAGTGTCTGGGCTGCCAAAGAACAGCAGTGCCCTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAAAGGCACCCACAGAACGCCAAACAAGCATTCCAGA
GCCTGCCAGCAATTCTCAAACAAATGTCAGCTAAGAACGCTTGCTCTGCCTTGTAGGAGCTCTG
AGCGCCCACACTTCCAATTAAACATTCTCAGCCAAGAACAGTGGACACACCTACCAAGACACTC
TTCTTCTCCCACCTCACTCTCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAGCA
TGTTTTCAAGATCATTGTTGTTGCTCTCTAGTGTCTTCTCTCGTCAGTCTTAGCCT
GTGCCCTCCCTACCCAGGCTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTTCCT
AGCTAGTGTCAATTAAACCTAAATGCAATCAGGAAAGTAGCAAACAGAACAGTCAATAAAATT
AAATGTCAAAAAAAAAAAAAAA

FIGURE 26

MKVLISSLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKQCPKCDHFKGKVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAGCCAGTGC~~CCC~~CAGCGGAAGCACAGCTCAG
AGCTGGTCTGCCATGGACATCCTGGTCCACTCCTGCAGCTGCTGGTCTGCTCTTACCC~~T~~GCC
CCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCGTGCAAAAGCTACTTCCC~~T~~ACCTGA
TGGCCGTGCTGACTCC~~A~~AGAGCAACCGAAGATGGAGAGCAAGAACGGGAGCTTCA~~G~~CCAG
ATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCC~~T~~ACTGGAGCTGGGCTGCGGAACCGAGC
CAACTTCAGTTCTACCCACCGGCTGCAGGGTACCTGCCTAGACCCAAATCCCCACTTGAGA
AGTTCTGACAAAGAGCATGGCTGAGAACAGGCCACCTCAATATGAGCGGTTTG~~G~~GGCTCCT
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGT~~G~~CTGTG
CTCTGTGCAGAGCCAAGGAAGGTCTGCAGGAGGTCCGGAGAGTACTGAGAC~~CCGGAGGT~~GTGC
TCTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGCCTCATGTGGCAGCAAGTTTC
GAGCCCAC~~T~~GGAAACACATTGGG~~G~~ATGGCTGCCTCACCAGAGAGAC~~T~~GAAGGATCTGA
GAACGCCAGTTCTCGAAATCCAATGGAACGACAGCCCCCTCC~~T~~GAAGTGGTACCTGTTG
GGCCCCACATCATGGGAAAGGCTGCAAACAAATCTTCCAAGCTCCAAGGC~~ACT~~CATTGCTCC
TTCCCCAGCCTCAATTAGAACAAAGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCTA
GCAGAACATGAGAGAACATT~~C~~ATGTACCACCTACTAGTCCCTCTCCCCAAC~~T~~CTGCCAGGGC
AATCTCTAACTTCAATCCGCC~~T~~CGACAGT~~G~~AAAAGCTCTACTTCTACGCTGACCCAGGGAGG
AAACACTAGGACCC~~T~~GTGTATCCTCAACTGCAAGTTCTGGACTAGTCTCC~~A~~ACGTTGCCTC
CCAATGTTGCC~~T~~TCCTCGTCC~~T~~CATGGTAAAGCTCCTCTCGCTTCC~~T~~GAGGCTACAC
CCATGCGTCTCTAGGA~~A~~CTGGTACAAAAGTCATGGTGCCTGCATCC~~T~~GCCAAGCCCCCTGAC
CCTCTCTCCCCACTACCACCTCTTCC~~T~~GAGCTGGGGCACCAGGGAGAATCAGAGATGCTGGGG
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTGTTCTCAAATATTTTTAATAAATAGACGAA
ACCACG

FIGURE 28

MDILVPLLQLLVLLTLPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSQIKGL
TGASGKVALLELGCCTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDVVVCTLVLCVQSPRKVLQEVRRVLPGVLFFWEHVAEPYGSWAFMWQQVFEPTW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPGPHIMGKAVKQSFPSKALICSFPSL
QLEQATHQPIYLPRLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTGCCTATCCACCTCCCCAAGCCCCTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTGCTAAAGGCTCATGCTGGAGTGGGACTGGTCGGTCCCAGAAAGTCTCTCTG
CCACTGACGCCCATCAGGGATTGGCCTCTTCCCCCTCCTGTGTCTCCTGCCTCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGAAGGGAGAAAGTGGGGATGGCTA
AGAAAGCTGGGAGATAGGAAACAGAAGAGGGTAGTGGGTGGCTAGGGGGCTGCCTTATTAAA
GTGGTTGTTATGATTCTTACTAATTACAAAGATATTAAGGCCCTGTTCATTAAGAAATT
GTTCCCTTCCCTGTGTTCAATGTTGTAAAGATTGTTCTGTGTAAATATGTCTTATAATAAAC
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MLLLTLLLLLKKGSCLEWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTCSQAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTTGAATTCCCTCAACTATACCCACAGTCCAAAGCAGACTCACTGTGTCCCAGGCTACCAGTT
CCTCCAAGCAAGTCATTCCTTATTTAACCGATGTGTCCTCAAACACCTGAGTGCTACTCCCT
ATTTGCATCTGTTTGTATAAATGATGTTGACACCCCTCCACCGAATTCTAAGTGAATCATGTCGG
GAAGAGATAACATCCTGGCCTGTGTATCCTCGCATTAGCCTGCTTGGCATGATGTTACC
TTCAGATTCACTACCACCCCTCTGGTCACATTTCTTCATTGTTATTTGGGATTGTTGTT
TGTCTGCGGTGTTTATGGTGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGGACA
CAGAAAGGAAAATATGAAGTGCCTGGGGTTGCTATCGTATCCACAGGCATCACGGCAGTG
CTGCTCGTCTGATTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTCCAAATCAC
AAATAAAGCCATCAGCAGTGCTCCCTCCTGCTGTTCCAGCCACTGTGGACATTGCCATCCTCA
TTTCTCTGGGCTCTGGTGGCTGTGCTGCTGAGGCTGGGAACTGCAGGAGCTGCCAGGTT
ATGGAAGGCGGCCAAGTGAATATAAGCCCTTCGGCATTGGTACATGTGGTGTACCTT
AATTGGCCTCATCTGGACTAGTGAATTCATCCTGCGTGCCAGCAAATGACTATAGCTGGGCAG
TGGTTACTGTTATTCACAGAAGTAAAAATGATCCTCCTGATCATCCCATCCTTCGTCTC
TCCATTCTCTTCTTCTACCATCAAGGAACCGTTGTGAAAGGGTCTTTAATCTGTGGTGT
GATTCCGAGAATCATTGTCATGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGT
CCAGGTACCTGTCGATGCTGCTACTGCTGTTCTGGTGTCTGACAAATACCTGCTCCATCTC
AACCAGAATGCATATACTACAACGTCTTACATCTATTAACTGCTTGGAGACTTCATAA
ATTCAAATCTTGTCCAAGAAACTCAAGTCACCTTACATCTATTAACTGCTTGGAGACTTCATAA
TTTTCTAGGAAAGGTGTTAGTGGTGTGTTCACTGTTGGAGGACTCATGGCTTTAACTAC
AATCGGGCATTCCAGGTGTGGCAGTCCCTCTGTTATGGTAGCTTTGCCTACTTAGTAGC
CCATAGTTTTATCTGTGTTGAAACTGTGCTGGATGCACCTTCCGTGTTGCTGTTGATC
TGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTATGGATCAAGAATTCTGAGTTCGTA
AAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGAATGAGGA
GGGAACAGAACTCCAGGCCATTGTGAGATGATACCCATTAGGTATCTGTACCTGGAAAACATT
TCCTTCTAAGAGCCATTACAGAAATAGAAGATGAGACCCTAGAGAAAAGTTAGTGAATTTTT
TTAAAAGACCTAATAACCCATTCTCCTCAAAA

FIGURE 32

MSGRDTILGLCILALALSLAMMFTFRFITLLVHIFISLVLGLLFVCGLWWLYYDYNLDSIE
LDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISAPFLLFQPLWTFA
ILIFFWVLWVAVLSSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQQMTIA
GAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVMVMQNALKEQQHG
ALSRYLFRCYCFCWCLDKYLLHLNQNAYTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD
IIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAHSFLSVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTCTGGACCTCTCCCTGTTCTCCTTAGA
ATAATTGTATGGGATTGTGATGCAGGAAGCTAAGGGAAAAAGAATATTCAATTCTGTGTGGT
GAAAATTGGAAAAAAATTGCCTCTCAACAGGGTGTCAATTCTGATATTTATGAGGAC
TGTTGTTCTCACTATGAAGGCATCTGTTATTGAAATGTCCTGTTGCTGGTACTGGAGTAC
ATTCAAACAAAGAACGGCAAAGAAGATTAAAGGCCAAGTCACTGTGCCTCAGATCAACTGC
GATGTCAAAGCCGGAAAGATCATGATCCTGAGTTCATGTGAAATGTCAGCAGGATGCCAAGA
CCCCAAATACCATGTTATGGCACTGACGTGATGCATCTACTCCAGTGTGTGGCCTGCCG
TACACAGTGGTGTGTTGATAATTCAAGGAGGAAAATACTGTCGGAAGGTTGCTGGACAGTCT
GGTTACAAAGGGAGTTATTCAACGGTGTCCAATGTTATCCCTACACGATGGAGAGAAATCCTT
TATCGTCTAGAAAAGTAAACCCAAAAGGGTGTAAACCTACCCATCAGCTCTTACATACATCAT
CGAAAAGTCCAGCTGCCAAGCAGGTGAGACCACAAAGCCTATCAGAGGCCACCTATTCCAGGG
ACAACGTGACAGCCGGTCACTCTGATGCACTGTTCTGGCTGTACTGTAGCTGTGGCCACCCCCAC
CACCTTCCAAGGCCATCCCCCTCTGCTGTTCTACCACAGCATCCCCAGACCAACATCAGTGG
GCCACAGGAGCCAGGAGATGGATCTCTGGTCCACTGCCACCTACACAAGCAGCCAAAAGGGCC
AGAGCTGATCCAGGTATCAAAGGCAAGATCCTTCAGGAGCTGCCCTCCAGAAACCTGTTGGAGC
GGATGTCAGCCTGGGACTTGTCCAAAAGAAGAATTGAGCACACAGTTGGAGGCCAGTATCCC
TGGGAGATCCAAACTGAAAATTGACTTGTGTTTAATTGATGGGAGCACCAGCATTGGCAA
CGCGGATTCCGAATCCAGAAGCAGCTCTGGCTGATGTTGCCAAGCTTGCACATTGGCCCTGC
CGGTCCACTGATGGGTGTTGTCAGTATGGAGACAACCTGCTACTCACTTAAACCTCAAGACAC
ACACGAATTCTCGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGAGACTTCTAAT
GTAGGTGGGCCATCTCTTGTGACCAAGAACTTCTTTCAAAGCCAATGGAAACAGAACGG
GGCTCCCAATGTGGTGGTGGTGTGGATGGTGGATGGCTGGCCACGGACAAAGTGGAGGGCTTCAA
GACTTGGAGAGAGTCAGGAATCAACATTCTCATCACCATTGAGGTGCTGCTGAAAATGAG
AAGCAGTATGTTGGTGGAGGCCAACTTGCAAAACAGGCCGTGTGCAAGAACAAACGGCTTCAACTC
GCTCCACGTGCAAGAGCTGGTTGGCCTCCAAAGACCCCTGCAGCCTCTGGTGAAGCGGGTCTGCG
ACACTGACCGCCCTGGCCTGCAGCAAGACCTGCTGAACTCGGCTGACATTGGCTCGTCATCGAC
GGCTCCAGCAGTGTGGGGACGGCAACTTCCGCACCGTCTCCAGTTGTGACCAACCTCACCAA
AGAGTTTGAGATTCCGACACGGACACGCGCATCGGGGGCGTGCAGTACACCTACGAACAGCGGC
TGGAGTTGGTTCCGACAAGTACAGCAGCAAGCCTGACATCTCAACGCCATCAAGAGGGTGGGC
TACTGGAGTGGTGGCACCAGCACGGGGCTGCCATCAACTTCGCCCTGGAGCAGCTTCAAGAA
GTCCAAGCCCAACAGAGGAAGTTAATGATCCTCATCACCAGCAGGGAGGTCTACGACGACGTCC
GGATCCCAGCCATGGCTGCCATCTGAAGGGAGTGTACACCTATGCGATAGGCGTTGCCCTGGCT
GCCCAAGAGGAGCTAGAAGTCATGCCACTCCCCGCCAGAGACCACTCCTCTTGTGGACGA
GTTTGACAAACCTCCATCAGTATGTCCTCAGGATCATCCAGAACATTGTCAGAGTTCAACTCAC
AGCCTCGGAACTTGAATTCAAGAGCAGGAGCAGCACCAGCAAGTGTCTTACTAATGACGTGTT
GGACCACCCACCGCTTAATGGGACGACGGTGCATCAAGTCTGGCAGGGCATGGAGAAC
AAATGTCCTGTTATTATTCTTGCATCATGCTTTTCAATTCCAAAATGAGGTTGAGAAC
TGATCACAAACGTATAGAATGAGCCAAAGGCTACATCATGTTGAGGGTGTGGAGATTTCAT
TTTGACAAATTGTTCAAAATAATGTCGAATACAGTGCAGCCCTACGACAGGCTTACGCTAG
AGCTTTGTGAGATTGTTAAGTGTATTCTGATTTGAACTCTGTAACCTCAGCAAGTTCAT
TTTGTCATGACAATGAGGAATTGCTGAATTAAATGTTAGAAGGATGAAAAA
AAA
AAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAG

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
CQDPKYHVGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWR
ESFIVLESKPKKGVTVPSALTYSSSKPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLAVTVAVA
TPTTLPRPSPSAASTTSIPRPQSVGHRSQEMDLWSTATYTSSQRPRADPGIQQRQDPSGAAFQKP
VGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI
GPAGPLMGVVQYGDNPATHFNLKHTNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVVMVWDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTQPLVKRVCDTRLACSKTCLNSADIGFVIDGSSSGTGNFRTVLQFVTN
LTKEFEISDTDTRIGAVQYTYEQRLEFGFDKYSSKPDILNAIKRVGWSSGTSTGAAINFALEQL
FKKS KPNKRKLMILITDGRSYDDVRI PAMA AHLKGVITYAIGVAWA AQEELEVIATHPARDHSFF
VDEF DNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCCTTAGGAGGTGGCTGCCTGTGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGCTTTTTCTGTTTCAGAGTAGTTACAACAGATCTGAGTGTTAATTAAGCATGGAAT
ACAGAAAACAACAAAAACTTAAGCTTAATTCTCATCTGGAATTCCACAGTTCTTAGCTCCCTGGACCC
GGTTGACCTGTTGGCTTCCCGCTGGCTGCTCTACACGTGGTCTCCGACTACTCACCCCGAGTGTA
AAGAACCTTCGCGTCGCTGCTTGAGCTGCTGTGGATGGCTCGGCTCTGGACTGTCCCTCCGAGTA
GGATGTCACTGAGATCCCTCAAATGGAGCCTCTGCTGTCACTCCTGAGTTCTTGATGTGGTAC
CTCAGCCTCCCCACTACAATGTGATAGAACCGGTAACTGGATGTACTCTATGAGTATGAGCGATTAA
CAGACAAGACTTCACCTCACACTCGAGAGCATTCAAACAGTCTCATCAAATCCATTCTGGTCATTC
TGGTGACCTCCCACCCCTCAGATGTGAAAGCCAGGCCATTAGAGTTACTTGGGTGAAAAAAAGTCT
TGGTGGGATATGAGGTTCTACATTTCTTATTAGGCCAAGAGGCTGAAAGGAAGACAAAATGTTGGC
ATTGTCCTAGAGGATGAAACACCTCTTATGGTACATAATCCGACAAGATTAGACACATATAATA
ACCTGACCTGAAAACCATTATGGCATTCAAGGTGGTAACCTGAGTTTGGCCAAATGCCAAGTACGTAATG
AAGACAGACACTGATGTTTCAATACTGCAATTAGTGAAGTATCTTAAACCTAAACCACTCAGA
GAAGTTTCACAGGTATCCTCTAATTGATAATTATTCTATAGAGGATTACCAAAAACCCATATT
CTTACCAAGGAGTATCCTTCAGGTGGTCCACTGCAGTGGGTGGTTATATAATGTCAGAGAT
TTGGTGCAAGGATCTATGAAATGATGGTCACGTAAAACCCATCAAGTTGAAGATGTTATGTCGGGAT
CTGTTGAATTATTAAAAGTGAACATTCAATTCCAGAAGACACAACTTTCTTCTATAGATCC
ATTGGATGTCGTCACAGACGTGTGATTGCAAGCCCAGGCTTCTCCAGGAGATCATCACTTT
TGGCAGGTCTGCTAAGGAACACCATGCCATTATTAACTCACATTCTACAAAAGCCTAGAAGGACAG
GATACTTGTGAAAGTGTAAATAAGTAGGACTGTGGAAAATTCAAGGGAGGTCACTGTGATTATTAGTCAGG
ACACTGAACCTGAAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTATTAGTCAGG
CCCTTCAAAGATGATATGTGGAGGAATTAAATATAAGGAATTGGAGGTTTGCTAAAGAAATTAGG
ACCAAAACAATTGGACATGTCATTCTGAGACTAGAATTCTAAAAGGTGTTACTGAGTTATAAGCTCA
CTAGGCTGAAAAACAAACATGTAGAGTTATTGAAACAATGTAGTCACCTGAAGGTTTGCTGA
TATCTTATGTGGATTACCAATTAAAAATATATGTAGTTCTGTGTCAAAAAACTCTTCACTGAAGTTATA
CTGAACAAATTTCACCTGTTGGTCATTATAAGTACTTCAGATGTTGCAGTATTTCACAGTTATT
ATTATTTAAATTACTTCACCTTGTTAAATGTTGACGATTCAACAAAGATAAAAGGATAG
TGAATCATTCTTACATGCAACATTTCCAGTTACTTAACGTGATCAGTTATTATTGATAACATCACTCCA
TTAATGTAAAGTCAGGTCAATTGCAATATCAGTAATCTCTGGACTTGTAAATTTACTGTGGT
AATATAGAGAAGAATTAAAGCAAGAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPIYRQDFHF
TLREHSNCSSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSSWWGYEVLTFFLLGQEAEKEDKMLA
LSLEDEHLLYGDIIHQDFLDTYNNLTLKTIIMAFRWVTEFCPNAKYVMKTDTDVFINTGNLVKYLL
NLNHSEKFFTGYPLIDNNSYRGFYQKTHISYQEYPDFKVFPPYCSDLGYIMSRLVPRYEMMGHV
KPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

FIGURE 38

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAEWNIMCRECCYEYDQIECVCPGKREVVGYT
IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGTLDDFYVKGFYCAECRAGWYGGDCMRCGQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDGQII
KRVCGNERPAPIQSISGSSLHVLFHSDGSKNFDGFHAIYEEITACSSPCFHDTGCVLDAKSYKC
ACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSGNE
KRTCQQNGEWSGKQPIKACREPCKISDLVRRRVLPQMVSRETPLHQLYSAAFSKQKLQSAPTK
KPALPFGDLPGMYQHLLHTQLQYECISPFYRRLGSSRRTCLRTGKWSGRAPSCIPCGKIENITAP
KTQGLLRWPWQAAIYRRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAHCVTDLGKVTMIKTADL
KVVLGKFYRDDDREKTIQSLQISAIILHPNYDPILLDADIAILKLLDKARISTRVQPICLAASR
DLSTSFOQESHTVAGWNVLADVRSPGFKNLTLRGVVSVVDSLLCEEQHEDHGIPVSVDNMFCA
SWEPTAPSIDIETGGIAAVSFPGGRASPEPRWHLMGLVSWSYDKTCSHRLSTAFTKVLPPFKDWI
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTTCCATCCTCTCATCTGAGAATCAGAGAGCATAATCTCTTACGGGCCGTGATTTATAACGTGGCTTAATC
TGAAGGTTCTCAGTCAAAATTCTTGTGATCTACTGATTGTGGGGCATGCCAAGGTTGCTTAAAGGAGCTTGGCTGG
TTTGGGCCCTGTAGCTGACAGAAGGTGCCAGGGAGAATGCAGCACACTGCTGGAGAATGAAGGCCTCTGTTGC
TGGCTTGGCTTGGCTCAGTCCTGCTAATCATTGACAATGTGGCAACTGCACTTCTGTATTCAAACCTGTGA
AAGGTGCCCTCCACTACGCCCTGACCAAAGATAGGAAGAGGCCGTACAAGATGGCTGCCAGACGGCTGTGCGAGCC
TCACAGGCCAGGCTCCCTCCCGAGGGTTCTGCAGCTGCCACCATCTCCTTAATGACAGACGAGCCTGGCTAGACA
ACCTGCTACGTGTCTCGCAGGGACGGCAGCCAGCAATCAGCCAGTGGACTCTGGCCGGAGCAACCGAAGTA
GGGCACGGCCCTTGAGAGATCCACTATTAGACAGATCATTAAAAAAATAAATCGAGCTTGAGTGTCTCGAA
GGACAAAGACGGGAGTGCAGTTGCCAACATGCCGACCAGGGAGGGAAATTCTGAAAACACCAGTGCCCTGAAG
TCTTCTCAAGGTTGTAACACCTGATTCCAGATGGTGAATTACAGCATCAAGATCAATCGAGTAGATCCAGTGA
GCCTCTCTATTAGGCTGGGGAGGTAGCAGAACCCACTGGCCATATCATTATCAAACACATTATCGTATGGGG
TGATCGCCAGAGACGCCGGCTACTGCCAGGAGACATCATTCTAAAGGTCAACGGATGGACATCAGCAATGTCCCTC
ACAACTACGCTGTGCGTCTCCTGCCAGGCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAATTCGCCA
GCAGGAACAATGGACAGGCCCGATGCCATCAGACCCAGATGACAGCTTCATGTGATTCTCAACAAAAGTAGCC
CCGAGGAGCAGCTGGATAAAAATGGTGCAGCAGGAGAATGACCGTGTGTTAGCCATCAATGGACATGATCTCGATATGGCA
GTGTGGCATATGACATGGTCACTGAGGAGAATGACCGTGTGTTAGCCATCAATGGACATGATCTCGATATGGCA
GCCAGAAACTGCCGCTCATCTGATTCAAGGCCAGTGAAGAGACGTGTCACCTCGTGTCCCCCAGGTCGGCAGC
GGAGCCCTGACATCTTCAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCCAGGGAGAGGAGCAACAA
CTCCCAAGCCCTCCATCTACAATTACTGTGATGAGAAGGTGTTAAATCTCAAAAGACCCGGTGAATCTCTCG
GCATGACCGTGCAGGGAGCATCACATAGAGAATGGGATTTGCCATCTATGTCATCAGTGTGAGCCGGAGGAG
TCATAAGCAGAGATGGAAGAATAAAAACAGGTGACATTGTTGAATGTGGATGGGTCGAACGTGAGGTCAGCC
GGAGTGAGGCGAGTGGCATTATTGAAAAAGAACATCATCTCGATGACTCAAGCTTGGAGTCAAGAGTATGAGC
CCCAGGAAGACTGCAGCAGGCCAGCAGCCCTGGACTCCAACACATGGCCCAACCCAGTGAATGGTCCCCATCCT
GGGTGATGTGGCTGGAATTACCCAGGTGCTGTATACTGTAAGAGATATTGTTACAGAAGAAACACAGCTGGAGTC
TGGGCTCTGCATTGTTAGGAGGTTATGAGAAGAACATCAATGGAAACAAACCTTTTCACTCAAATCATTGTTGAAGGAA
CACCAGCATACATGATGGAAGAATTAGATGTGGTGAATTCTCTGCTGCAATGGTAGAAGTACATCAGGAATGA
TACATGCTTGGCAAGACTGCTGAAAGAACCTAAAGGAGAACATTACTCTAACATTGTTCTGGCTGGCACT
TTTATAGAATCAATGATGGTCAAGAGAACAGAAAAATCACAAATAGGCTAAGAAGTGAACACTATATTATC
TTGTCAGTTTATATTAAAGAAGAACATTGTAAGGAGGATGTCAGGAAAGTATGATCATCTAATGAAAGCCAGTT
ACACCTCAGAAAATATGATCCAAAAAATTAAGAACATTGTAAGGAGGATGTCAGGAAAGTATGATCATCTAATGAAAGCCAGTT
AACATTGTTATATTCTATTCAATAAAAAGCCAAAACAACATTGATTGATTTGATACCCACTGAATT
CAAGCTGATTTAAATTTAAATGGTATGCTGAGTCTGCCAAGGGTACATTATGCCATTAAATACAGCT
AAAAATTTTAAATGCTGAGAACAGTGTGAGGAAACAGTTGCTGTTCATCAAACAAAGAATTAATTTTCAAGGTTAA

FIGURE 40

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDGASLTATAPS
PEVSAATISLMTDEPGLDNPAYSSAEDGQPAISPVDGRSNRTRARPFERSTIRSRSFKKINR
ALSVLRRTKSGAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINRVDPSESLSIRLV
GGSETPLVHIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVM
REQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGVFI FMVLDGGVAYRHG
QLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRQSPDI FQEAGWNSNGSWSPG
PGERSNTPKPLHPTITCHEKVNNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVISRDGR
IKTGDIILNVDGVELTEVRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP
SDWSPSWVMWLELPRLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG
RIRCGDILLAVNGRSTSGMIHACLRLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTGAAGCTTT
CTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTCACGTAATAAAAACATGGGCTTCAACCTGACT
TTCCACCTTCCTACAAATTCCGATTACTGTTGCTGTTGACTTTGTGCCCTGACAGTGGTGGGGTGGC
CACCAAGTAACTACTCGTGGGTGCCATTCAAGAGATTCTAAAGCAAAGGAGTTCATGGCTAATTCC
ATAAGACCCCTCATTTGGGAAAGGGAAAAACTCTGACTAATGAACCACCAAGAAGGTAGAACTT
GACAACGTCCCTCTGTCTCCTACCTCAGAGGCCAGAGCAAGCTCATTTCAAACCAAGATCTCAC
TTTGGAAAGAGGTACAGGCAGAAAATCCCAAAGTGTCCAGAGGCCGTATGCCCTCAGGAATGTAAG
CTTTACAGAGGGTCGCCATCTCGTCCCCACCGGAACAGAGAGAAAACACCTGATGTACCTGCTGGAA
CATCTGCATCCCTTCCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCACTCCACCAGGCTGAAGG
TAAAAAGTTAATCGAGCAAACCTCTGAATGTGGGCTATCTAGAACCCCTCAAGGAAGAAAATTGGG
ACTGCTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTACAAGTGTGAGGAG
CATCCCAAGCATTGGTGGTGGCAGGAACAGCACTGGTACAGGTACAGTTACAGTGGATATTGG
GGGTGTACTGCCCTAACAGAGAGCAGTTTCAAGGTGAATGGATTCTCTAACAAACTACTGGGAT
GGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTCCCGGCCCTG
CCTGAAGTGGTAAATATAATGGTCTTCCACACTAGAGACAAAGGAATGAGGTGAACGCAGAACG
GATGAAGCTCTTACACCAAGTGTACAGACTGGAGAACAGATGGTGTAGTAGTTGTTCTTATAAAT
TAGTATCTGTGGAACACAATCCTTATATATAACATCACAGTGGATTCTGGTTGGTGCATGACCC
TGGATCTTGGTGTATGGAAAGAACTGATTCTTGTGCAATAATTGGCTAGAGACTTCAA
ATAGTAGCACACATTAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTCCTTTGTATTTCT
TAGCAGAGCTCCTGGTGTAGTAGTATAAAACAGTTGTAACAAGACAGCTTCTTAGTCATTGAT
CATGAGGGTAAATATTGTAATATGGATACTGAAAGGACTTTATATAAAAGGATGACTCAAAGGATAA
AATGAACGCTATTGAGGACTCTGGTGAAGGAGATTATTAAATTTGAAGTAATATATTATGGGAT
AAAAGGCCACAGGAAATAAGACTGCTGAATGCTGAGAGAACCCAGAGTTCTCGTCCAAGGTAGAA
AGGTACGAAGATAACAATACTGTTATTCACTTATCCTGTACAATCATCTGTGAAGTGGTGGTGTAGGT
GAGAAGGCGTCCACAAAGAGGGAGAAAAGGCGACGAATCAGGACACAGTGAACCTGGGAATGAAGA
GGTAGCAGGAGGGTGGAGTGTGGCTGCAAAGGCAGCAGTAGCTGAGCTGGTGCAGGTGCTGATAGC
CTTCAGGGGAGGACCTGCCAGGTATGCCCTCAGTGTAGGCCACCAAGAGAATACATTCTTATTAGT
TTTAAAGAGTTTGTAAATGATTTGTACAAGTAGGATATGAATTAGCAGTTACAAGTTACAT
ATTAACATAATAATATGTCTATCAAATACCTCTGTAGTAAATGTGAAAAGCAAA

FIGURE 42

MGFNLTFHLSYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQOLDYGIYVIHQAEGKKFNRAKLLNVGYLEALKEENWDCFIFHDV
DLVPENDFNLYKCEEHPKHLVGRNSTGYRLRYSGYFGGVTAALSREQFFKVNGFSNNYWGWGGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGATGGACTGAGCTCCAGATCTGGG
CCGCTTGCCTCCTGCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTCCCACAAACAG
ACGGGACAACCTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCAGCTGGATGCCAT
GTTCCAGAGGCAGAGGAGGCAGACACCCACTTCCCCATCTGCATTCTGCTGCGGCTGCTGTC
ATCGATCAAAGTGTGGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCCGTCCCCTCCC
TTCCTTATTATTCTGCTGCCAGAACATAGGTCTTGAATAAAATGGCTGGTTCTTTGTTT
TCCAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 44

MALSSQIWAACLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRDTHFPI
CIFCCGCCHRSKCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTCCCCAACATGCCTCACCC
TCATCTATATCCTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT
TCCGTTGGTGGGGCCGTGACTTCCCCCTGAAGTCAAAGTAAAGCAAGTTGACTCTATTGTCTG
GACCTTCAACACAACCCCTCTTGTACCACAGCCAGAAGGGGGACTATCATAGTGACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACGTGAAG
AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCACCCA
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGAAGAGGATGTGATT
TATAACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTC
CTGGAGATGGGGAGAAAGTGAATATGACCTTCATCTGCCAGGAACCTGTCAAGCAGAAACT
TCTCAAGCCCCATCCTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCCCTCCATG
GTCCTCCTGTGTCTCTGTTGGTGCCTCCTGCTCAGTCTCTTGACTGGGTCTATTCTTG
GTTTCTGAAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAGAGTGGACATTGTCGGG
AAACTCCTAACATATGCCCTCATCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAACTCTAAAGGAAGATCCAGCAAATACGGTTACTCCACTGTGGAAATACGAAAAAGAT
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTCCTATGAGAATGTTA
TCTAGACAGCAGTGCACTCCCCTAAGTCTGCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSGAVTFPLSKVKQVDSIVWTFTPLVTIQP
EGGTIIVTQNRNRERVDFPDGGYSLKLSKIKKNDSGIYYVGIYSSSLQQPSTQEVVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLILFVLGLFLWFLKRERQEEYIE
EKKRVDICRETPNICPHSGENTEYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTCAGCCTGCTGGTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATAACCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTTCTCAAAACCCATCTCTGCTTGAGTGGTGGTCCCAGGAATT
ATAGGAGCAGGTCTGATGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTG
CTGCAACAAACAGAACTGGAATGTTCTTCATCATTTTCAGTGTGATCACAGTCATTGGTGCCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTCTAAAGGTCCCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCATTGTGAATTTCATTGAAAAACATCAGTGACATTCCAGAATCCTCAA
CTTGCAGTGGTTTCAATGACTCTTGTGCACCTCCTACTGGTTCAATAAACCCACCAAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTCCACTTCGATTCTGAAGAAAACAAACATAGG
CTTATCCACTCTCAGTATTTAGGTCTATTGCTTGGATTCTGGAGGTCCCTGGGGCT
CAGTCAGATAGTCATCGGTTCTGGCTGTGTGGAGTCTAAGCGAAGAAGTCAAATTG
TGTAGTTAATGGGAATAAAATGTAAGTATCAGTAGTTGAAAAAAAA

FIGURE 48

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNANCEFSL
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSHFDEENKHRLIHFSVFLGL
LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCGCGCTGCCAGCTCAGGTGAGCCCTGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGGCTGCAGACCCCCGCCAGTGCCTCTCCCCCTGCAGCCCTGCCCTGAACTGTGA
CATGGAGAGAGTGACCCCTGGCCCTCTCCTACTGGCAGGCCTGACTGCCTTGGAAAGCCAATGACC
CATTTGCCAATAAAGACGATCCCTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGC GGAGGGCTCCTGGCCATTGCTGGGATCGCGGAGTCTGAGTGGCAAATGCAAATACAAGAG
CAGCCAGAAGCAGCACAGTCCGTACCTGAGAAGGCCATCCCACTCATCACTCCAGGCTCTGCCA
CTACTTGCTTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCCAGCACC
TCCTCCCTGGGAGGCCTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCCCT
TTCTGATCAGGAGGCTTCTTATGAATTAAACTCGCCCCACCACCCCTCA

FIGURE 50

MERVTLALLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGLLAIAGIAAVLSGKCKYKS
SQKQHSPVPEKAIPPLITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
AGGACAGGGAGTCGGAAGGAGGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG
AGGAGACCCCTGGTGGAGGAAGACACTCTGGAGAGAGAGGGGCTGGCAGAGATGAAGTTCCAG
GGGCCCCCTGGCCTGCCTCCTGCTGGCCCTGCTGGCAGTGGGAGGCTGGCCCTGCAGAG
CGGAGAGGAAAGCACTGGGACAAAATTGGGGAGGCCCTGGACATGGCCTGGAGACGCCCTGA
GCGAAGGGGTGGAAAGGCCATTGGCAAAGAGGCCGGAGGGCAGCTGGCTCTAAAGTCAGTGAG
GCCCTGGCCAAGGGACCAGAGAACAGCTGGCACTGGAGTCAGGCAGGTTCCAGGCTTGGCAG
AGCAGATGCTTGGCAACAGGTCGGGAAGCAGCCATGCTCTGGAAACACTGGCAGGAGA
TTGGCAGACAGGAGAAGATGTCATTGACACGGAGCAGATGCTGTCCGCGCTCCTGGCAGGGG
GTGCCTGGCCACAGTGGTGGCTTGGAAACTCTGGAGGCCATGGCATTTGGCTCTCAAGGTGG
CCTTGGAGGCCAGGGCCAGGGCAATCCTGGAGGTCTGGGACTCCGTGGTCCACGGATAACCCG
GAAACTCAGCAGGCAGCTTGGATGAATCCTCAGGGAGCTCCCTGGGTCAAGGAGGCAATGG
GGGCCACCAAACCTTGGACCAACACTCAGGGAGCTGTGGCCAGCCTGGCTATGGTCAGTGAG
AGCCAGCAACCAGAATGAAGGGTGCAGCAATCCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA
ACTCTGGGGAGGCAGCGGCTCACAGTCGGCAGCAGTGGCAGTGGCAGCAATGGTACAACAAC
AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGTGGCAGCAG
TGGCGGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGAGT
CCTCCTGGGATCCAGCACCGGCTCCTCCGGCAACCACGGTGGAGCGGGAGGAAATGG
CATAAACCCGGGTGTGAAAAGCCAGGGATGAAGCCGGAGCGGGAAATCTGGATTCAAGGG
CTTCAGAGGACAGGGAGTTCCAGCAACATGAGGGAAATAAGCAAAGGGCAATGCCCTCTT
GAGGCTCTGGAGACAATTATCGGGGCAAGGGTCAGCTGGGAGCTGGAGGAGGTGACGCTGTT
GGTGGAGTCATACTGTGAACTCTGAGACGTCCTGGATGTTAACCTTGACACTTCTGGAA
GAATTTAAATCCAAGCTGGTTCATCAACTGGGATGCCATAAACAAAGGACCAAGAGCTCTC
GCATCCCGTGACCTCCAGACAAGGAGCCACAGATTGGATGGAGCCCCACACTCCCTCTAA
AACACCACCCCTCTCATCACTAAATCTCAGCCCTGCCCTGAAATAACCTTAGCTGCCCCACAAA
AA
AA

FIGURE 52

MKFQGPLACLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGAAGSKVS
EALGQGTREAVGTGVRQVPGFAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGVP
GHSGAWETSGGHGIFGSQGGLGGQGQGNPGGLGTPWVHGYPGNSAGSGFMNPQGAPWGQGGNGGPPNF
GTNTQGAVAQPGYGSVRASNQNEGCTNPPPSGSGGSSNSGGSGSQSGSSGSGSNGDNNNGSSSGS
SSGSSSSGSSGGSSGGSSGNSGGSRGDSGSESSWGSSTGSSGNHGGSGGNHKGCEKPGNE
ARGSGESGIQGFRGQGVSSNMREISKEGNRLGGSGDNYRGQGSSWGSGGGDAVGGVNTVNSETSPGM
FNFDTWFKNFKSKLGFINWDAINKDQRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGGACAAGCTGCTCCGACAGAAGGGATGTCGCTGCTGAGCCTGCCCTGG
CTGGGCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCCTGCTGCTGGTTGTCGGCTCCTGGCT
ACTCGCCCGCATCCTGGCTTGACCTATGCCCTATAACAACTGCCGCCGGCTCCAGTGTTC
CACAGCCCCAAAACGGAACCTGGTTGGGTACCTGGGCTGATCACTCCTACAGAGGAGGGC
TTGAAGGACTCGACCCAGATGTCGCCACCTATTCCCAGGGCTTACGGTATGGCTGGTCCC
CATCCCCTTCATCGTTTATGCCACCCTGACACCATCCGGTCTATACCAATGCCCTAGCTGCC
TTGCACCCAAGGATAATCTCTTCATCAGGTTCTGAAGGCCCTGGCTGGAGAAGGGATACTGCTG
AGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGACGCCCTTCATTCAACATCCT
GAAGTCCTATATAACGATCTCAACAAGAGTGCACAAACATCATGCTGACAAGTGGCAGCACCTGG
CCTCAGAGGGCAGCAGTCGCTGGACATGTTGAGCACATCAGCCTCATGACCTGGACAGTCTA
CAGAAATGCATCTTCAGCTTGACAGCATTGTCAGGAGAGGCCAGTGAATATATTGCCACCAT
CTTGGAGCTCAGTGCCCTGTAGAGAAAAGAAGCCAGCATACTCCAGCACATGGACTTCTGT
ATTACCTCTCCCATGACGGCGCGCTTCCACAGGGCTGCCGCTGGTCATGACTTCACAGAC
GCTGTCATCCGGGAGCGCGCTCGCACCCCTCCACTCAGGGTATTGATGATTTTCAAAGACAA
AGCCAAGTCCAAGACTTGGATTTCATTGATGTGCTCTGCTGAGCAAGGATGAAGATGGAG
CATTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTGGAGGCCATGACACCACG
GCCAGTGGCCTCTCCTGGGTCTGTACAACCTTGCAGGGCACCCAGAAATACCAGGAGCGCTGCC
ACAGGAGGTGCAAGAGCTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGCC
AGCTGCCCTTCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCAGCTCCCTCATC
TCCCAGTGCACCCAGGACATTGTTCTCCAGATGGCCAGTCATCCCCAAAGGCATTACCTG
CCTCATCGATATTATAGGGTCCATCACAACCCAACTGTGTCGGCCGGATCCTGAGGTCTACGACC
CCTTCCGCTTGTACCCAGAGAACAGCAAGGGGAGGTACCTCTGGCTTTATTCCCTTCTCCGCA
GGGCCAGGAACGTGACATGGGCAGGCAGTCGCCATGGCGGAGATGAAAGTGGCCTGGCGTTGAT
GCTGCTGCACTTCCGGTCTGCCAGACCAACTGAGCCCCGAGGAAGCTGGAATTGATCATGC
GCGCCGAGGGCGGGCTTGGCTGCCAGGGCTGAATGTAGGCTGCAGATGACTTCTGAC
CCATCCACCTGTTTTGCAGATTGTCATGAATAAACGGTGCTGTCAAA

FIGURE 54

MSLLSLPWLGRLPVAMSPWLLLLVVGSWLLARILAWTYAFYNNCRLQCFQPPKRNWFWGHLG
LITPTEEGLKDSTQMSATYSQGFTVWLGPPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
WLGEGLLSSGGDKWSRHRMLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI
SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSDGRRFHRAC
RLVHDFTDAVIRERRTLPTQGIDDFFKDKAKSKTLDFIDVLLSKDEDGKALSDEDIRAEADTF
MFGGHDTTASGLSWVLYNLRHPEYQERCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPPEVYDPFRFPENSKGRSP
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN
VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAACAGCTGAAGCGAATGTTGA
GCCTACTCGTTGATTGCAACTATCATGGTGCTGTTGCTGCACTTACCCCTGTGTTCTGCCT
TTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTGCAGTCTTGGCATTGACGTGG
TACAGCCTTCCTTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTGTTGCCGTGTCTTGC
ATAATTCATGGCCAGTTTATGAAGCTTGGAAAGGCACATGGACAGAAGCTGGTGGACAGTTT
GTAACATCTCGAACACCTCTGTCTTACAGACATGTGCCCTTATCTTGCAGCAATGTGTTGCTT
GTGATTGAAACATTGAGGGTTACTTTGGAAGCAACAATACATTCTGAACCTGAATGTCAGTA
GCACAGGGATGAGAAGTGGGTTCTGTATCTTGTGGAGTGGAAATCTTCCTCATGTACCTGTTCTC
TCTGGATGTTGTCCCAGTGAATTCCCATGAATAACAAACCTATTCAAGCAACAGCAAAAAAAAAAAAA
AA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCASFWWHNKGLALIFCILQSLALTWYSLSFIPFAR
DAVKKCFAVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGAGTGGAGCACCCAGCAGGCCAACATGCTCTGTCGTGCCTG
TACGTGCCGGTCATCGGGGAAGCCCAAGCCGACTTCAACTTTCAGTCAAGGGCTCCCTGCCAGCTGAAGCT
ATTTCAAGCTCAGTGTCTTCATCCCTCCAGGATTCTCCACCTACCGCCAGTGGAAAGCAGAAAATTGTACAAGCT
GGAGATAAGGACCTTGATGGCAGCTAGACTTTGAAGAAATTGTCCATTATCTCAAGATCATGAGAAGAAGCTGAGG
CTGGTGTAAAGTTGGACAAAAAAGAATGATGGACGCATTGACGCGCAGGAGATCATGAGTCCCTGCCGGACTTG
GGAGTCAGATATCTGAACAGCAGGAGAAAAATTCTCAAGAGCATGGAAAAACGGCACGATGACCATCGACTGG
AACGAGTGGAGAGACTACCACTCTCCACCCGGAAAACATCCCGAGATCATCCTCTACTGGAAGCATTCCACG
ATCTTGATGTGGGTGAGAATCTAACGGTCCCGATGAGTTCACAGTGGAGGAGAGGAGACGGGGATGTTGGAGA
CACCTGGTGGCAGGAGGTGGGCAGGGCCGATTCAGAACCTGACGGCCCCCTGAGCAGGCTCAAGGTGCTCATG
CAGGTCCATGCCCTCCAGCAACACATGGCATTGTTGGCTACTCAGATGATTGAGAAGGAGGGGCCAGG
TCACTCTGGCCGGCAATGGCATCAAGCTCCCAAATGGCCCAAGGCTATGAGCATTCAAATTGAGCAG
ATCAAGGCCCTTGTGGTACTGACCAAGGAGACTCTGAGGATTCAAGGAGGGCTTGTGGCAGGGTCTTGGCAGGGCC
ATCGCCAGAGCAGCATCTACCAATGGAGGCTCTGAAGACCCGATGGCGCTGGGAAGACAGGCCAGTACTCAGGA
ATGCTGAGCTGCGCAGGAGGATCTGGCAGAGAGGGGGTGGCCCTTCTACAAAGGCTATGCTCCCAACATGCTG
GGCATCATCCCTATGCCGCATCGACCTGAGCTCACAGGAGGCTCAAGATGCTGGCTGAGCAGTATGAGCTG
AACAGCCGGACCCGGCTGTGGCTCCGGCACCAGTCCAGTACCTGTGGCAGCTGGCAGCTGGCAGCTAC
CCCCCTGGCCCTAGTCAGGACCCGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
TCAGG
GCTGTGAGCATCAGTACGTTGCTACAGAGAACCTGAAGATCACCTGGCGTGCAGTCGGCGTGACGGGGGGAGGG
CGCCCGCAGTGGACTCGCTGATCTGGCCAGGCTGGGTGTCAGGATCTCATCTGTGAAATGTCACAAACT
AAGCTGTCTGAGCCAGGAGCTGTGAAACCCCTAGACGACCCGGCAGGGAGGGAGGAGGAGGAGGAGGAGG
GTCCCTGCTGACCCAGCAGACCCCTCTGGTGTTCACGGAGAACCCAGGATTCCTAGGGTCCAGGTCAGCAGG
CTCCGGCTCACATGTGAAGGACAGGACATTTCAGTGCCTGCAATAGTGACGTTGGACCCCTGGAGGCCGCT
TAGTTCTCCATTTCACCCCTGCAGGAGCTGTGGCCACGGCCCTGCCCTGGCTGCCGTGCATCTCCCTGTGC
CCTCTGCTGCTGCCCTGCTGCTGAGGTAAGGTGGGAGGGGCTACAGCCACATCCCACCCCTGTCACATCCC
ATAATCCATGATGAAAGGGTGGAGGTCACTGGCTCCAGGCTGAGCTCCAAACCTACAGCATTGACGCCAACCTGG
TGTGAAGGAAGGAGGAAGGAGTCTGGCTGTGGTCACTGGCATCTGAGCCCTGCTGATGGCTGGGCTCTGGGATG
CTTGGGAGTGCAGGGGCTCGGGCTGGCTGGCTGCACAGAGGCAAGTGCCTGGGCTCATGGCTCTGAGCT
GGCCTGACCCCTGTCAGGATGGGCCACCTCAGAACCAAACACTCAGTGTCCCACCTGTCAGGGCAGTGGAGCA
CCATGTTGAGGGCAAGGGCAGAGGCTTGTGTTCTGGGGAGGAAAGGAAAGGAGGAGGAGGAGGAGGAGG
ACTGTTGGGAAAGGGTTTGTCCAGAAGGACAAGCCGGACAAATGAGCAGCTCTGTGCTTCCAGAGGAAGCAGG
GAGCAGGAGCTGGCTGACTGCTCAAGGTCTGGCTGACGCCCTGGGGCTCTGTCAACCCAGCAGGGCCGAGC
GGGACCAAGCCACATCCACTTGTCACTGCTGGAACCTATTATTTGATTTATTTGAACAGAGTTATGCT
AACTATTATGATGTTAAATTATAGCTGTCATTTCAGTCTGCAATTCTTATCATATTATGTTATGTTATGTT
GATTGTACCTCCCAAGCCGGCAACTGGGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
GTCCAGAGAAATTCTTGGGACTGGAGGGCAGAAAAGCCGGCAGAAGGAGGAGGAGGAGGAGGAGGAGG
GTTGGGAGGGCTGGCCCTAGGATTCTGGGTTGACTGGGGCTGGAGAGAGAGGGAGGAGGAGGAGGAGG
AACCTTGAGGTGAAATCCAGTTATTCTCGCTGCGAGGGTTCTTATTACTCTTCTGAATGTCAGGCCAG
TGAGGTGCTCTCACTGTAATTGTGGTGGCGGGGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
CTTCTGCTGCCCTTGCTTAACAATGCCGGCAACTGGGACCTCACGGTTGCACTTCCATTCCACAGAATGACCTGA
TGAGGAAATCTCAATAGGATGCAAGATCAATGCAAAATTGTTATATGAAACATATACTGGAGTCGTCAAAAG
CAAATTAGAAGAATTGGACGTTAGAAGGTTGTCATTAAAGCACCTTCTAATAAAAGTGTGAAAGCTGAAAAAA
AA

FIGURE 58

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKLDG
QLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEKILKSMKDNG
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTEERQTGMWRHLVAGGG
AGAVSRTCTAPLDRLKVLMQVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIK
FMAYEQIKRLVGSQETLRIHERLVAGSLAGAIAQSSIYPMEVLKTRMALRKTGQYSGMLDCARR
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNAWLQHYAVNSADPGVFVLLACGTMSSTC
GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGCAGCTCCACTCAGCCAGTACCCAGATAACGCTGGAACCTTCCCCAGCCATGGC
TTCCCTGGGGCAGATCCTCTTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGTTGGTATTCAGGGAGACACTCCATCACAGTCACTACTGTCGCCTCAGCT
GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTGAACCTGACATCAAACCTTCTGATAT
CGTGATACATGGCTGAAGGAAGGTGTTAGGCTTGGTCCATGAGTTCAAAGAAGGCAAAGATG
AGCTGTCGGAGCAGGATGAAATGTTAGGCTTGGTCCATGAGTTCAAAGAAGGCAAAGATG
GGCAATGCCTCTTGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAAACTGGAGCCTTCAGCATGCCGG
AACTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCCTGTGAGGCTCCCCGATGGTCCCC
CAGCCCACAGTGGCTGGGATCCCAAGTTGACCAGGGAGCCAACCTCTCGGAAGTCTCCAATAC
CAGCTTGAGCTGAACCTCTGAGAATGTGACCATGAAGGTTGTCTGTGCTCTACAATGTTACGA
TCAACAAACACATACTCCTGTATGATTGAAAATGACATTGCCAAGCAACAGGGATATCAAAGTG
ACAGAACATGGAGATCAAAGGGAGTCACCTACAGCTGCTAAACTCAAAGGCTCTGTGTGT
CTCTCTTCTTGCCATCAGCTGGCACTTCTGCCTCTCAGCCCTTACCTGATGCTAATAAT
GTGCCTTGGCCACAAAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTCAC
CACCAGATATGACCTAGTTTATATTCTGGGAGGAATGAATTCATATCTAGAAGTCTGGAGTG
AGCAAACAAGAGCAAGAAACAAAAAGCAAGCCAAAGCAGAAGGCTCCAATATGAACAAGATAAAAT
CTATCTCAAAGACATATTAGAAGTGGAAAATAATTCTGTAACAGTGTGAACTAGACAAGTGTGTTAAGA
GTGATAAGTAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGT
CACCTGGGAGTGGAGGACAGGGATAGTGCATGTTCTGTGAATTGGTATGTTAGTTATGTGC
TGTAATGTTGCTCTGAGGAAGCCCCTGGAAAGTCTATCCAAACATATCCACATCTTATATTCCAC
AAATTAAAGCTGTAGTATGTACCTAACGACGCTGCTAAATTGACTGCCACTCGCAACTCAGGGCG
GCTGCATTTAGTAATGGGCTAAATGATTCACTTTTATGATGCTTCAAAGGTGCCTGGCTTC
TCTTCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTAGCATAAACAGAGCAGT
CGGGGACACCGATTTATAAATAAAACTGAGCACCTTCTTTAAACAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 60

MASLGQILFWSIISIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVGLVHFKEGKDELSSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAGTYKC
YIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVVSVLYNVTINNTYSCMIENDIAKATGDIKTESEIKRRSHLQLLNSKASL
CVSSFFAISWALLPLSPYLMLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCAGAATCACCATGGCAGCTATCCTTACCGGAGGGCTGCCAGGAGCTGCAGGACAAG
CACCAAGGAGCCCTCCGGTAGCTACTACCCCTGGACCCCCAATAGTGGAGGGCAGTATGGTAGT
GGCTACCCCTGGTGGTTATGGGGCTCTGCCCTGGAGGGCTTATGGACCACCAGCTGG
TGGAGGGCCATGGACACCCAATCCTGGATGTTCCCTCTGAACCTCCAGGAGGACCATATG
GCGGTGCAGCTCCGGGGCCCTATGGTCAAGCTACGGTCCCAGCAGCCT
GGCTTATGGACAGGGTGGCCCTCCAATGTGGATCCTGAGGCCTACTCCTGGTCCAGTC
GGTGGACTCAGATCACAGTGGCTATATCTCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTGACAAGACCAAGTCA
GGCCGCATCGATGTCACGGCTCTCAGCCCTGTGGAAATTCACTCCAGCAGTGGAAAGAACCTTT
CCAGCAGTATGACCGGGACCGCTGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGCCC
AAATGGCTACAACCTGAGCCCCAGTTCACCCAGCTCTGGTCTCCGCTACTGCCACGCTCT
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA
GGCCTCCGGAGAAGGACACAGCTGACAAGGCAACATCCGGCTCAGCTCGAGGACTTCGTCA
CCATGACAGCTCTCGGATGCTTGACCAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT
TCCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACATCTCTTCTTCTGTCCCTAGAAGAAC
ATTCTCCCTTGCTTGATGCAACACTGTTCAAAAGAGGGTGGAGAGTCTGCATCATGCCACCA
AAATGTGAGGACCGGGCTGAGGCCACACAGATAGGGCCTGATGGAGGAGAGGATAGAAGTTGA
ATGTCTGATGCCATGAGCAGTTGAGTGGCACAGCCTGGCACAGGAGCAGGTCTTGTAATGG
AGTTAGTGTCCAGTCAGCTGAGCTCCACCCGTATGCCAGTGGTAGTGTTCATGCCCTGTAC
GTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCACAGGCCATTCTCCAAAGTGGAAAT
CTGACCAAGCATGAGAGAGATCTGCTATGGGACCAAGTGGCTGGATTCTGCCACACCCATAAAAT
CCTTGTGTGTTAACTCTAGCTGCCCTGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGCAT
CTTGCCAGGCTCTGCCCTGCAGCTGGACCCCTCACTTGCCCTGCCATGCTCTGCTCGGCT
TCAGTCTCCAGGAGACAGTGGTCACCTCCCTGCCAATACTTTTTAATTGCATTTTTTC
ATTTGGGCCAAAGTCCAGTGAAATTGTAAGCTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSPGTPGGPYGGAAPGGPYGQPPPSSYGAQQPGLYGQGGAPPNVDPEAYSWFQSVDSDH
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTKSGRIDVYGFSALWKFIQQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSQFQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTGGGCCTGCTCCTGGCTGTCTTCATC
TCCCAGGCCTCTTGCCCGAGCATCGTGTGGAGGAGAAAGTTCCAAAATTCGGGACC
AACTTGCTCAGCTCGGACAACCTTCCTCCACTGGCCCTCTAACCTGAACATCCGCAGCCGC
TCTGGACCCTAGGTCTAATGACTTGGCAAGGGTCTCTGAAGCTCAGCGTGCTCCATCAGATG
GCTTCCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGCCCTCATCGTGGGGCTGCCTGCCATG
GATTCTGGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCCGTGAGGACGCCCTGGGG
AGCGCTGCCTGAAGAACTCTCTTACCTCTCCAGTGCTGCCCTCGCTCCGGCAGTGGCCCTT
TGCCTGGGAGTCTCTCCGATGCCACAGGCCTCTCACCTGAGGCTCACTCCACCAAGGAC
TCGGAGTCCAGGACTGCCCGTTCTAACCTGAGGCGGGGAAAATCTTCCAAACG
CCCTCCCTGGTCTCTCATCCACAGGGTTCTGCCTGATCACCCCTGGGTACCCCTGAATCCCAGTG
TGTCTGGGAGGTGGAGGCCCTGGACTGGTGGAAACGAGGCCATGCCACACCTGAGGG
ATCTGGGTATCAATAATCAACCCCCAGGTACCAAGCTGGGAAATATTAATCGGTATCCAGGAG
CAGCTGGGAAATATTAATCGTATCCAGGAGCAGCTGGGAAATATTAATCGGTATCCAGGAG
GCAGCTGGGAAATATTCATCTATACCCAGGTATCAATAACCCATTCTCTGGAGTTCTCCGC
CCTCCTGGCTCTCTTGGAACATCCCAGCTGGCTCCCTAATCTCCAAGCCCTAGGTTGCAGTG
GGGCTAGAGCACGATAGAGGGAAACCAACATTGGGAGTTAGAGTCCTGCTCCGCCCTTGCTG
TGTGGCTCAATCCAGGCCCTGTTAACATGTTCCAGCACTATCCCCACTTTCACTGCCTCCCC
TGCTCATCTCCAATAAAATAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

FIGURE 64

MQGRVAGSCAPLGLLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGPSNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPAGGSAVQRWPPSWGLPAMDSWPPEDPWQMMAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRSNSLGAGGKILSQRP
PWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRPMPHPEGIWGINNQPPGTWGNINRYPGGS
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPGVLRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCATCCCAGGAGGCCAGTGGCCACTATGGGTC
TGGGCTGCCCTTGTCCCTCTTGCACCTCTGGCAGCTCACATGGAACAGGGCCGGGTATGA
CTTGCAACTGAAGCTGAAGGAGTCTTCTGACAAATTCCCTCATGAGTCCAGCTTCCGGAA
TTGCTTGAAGCTCTGCCTCCCTCCATCTCCCTCAGGGACCAGCGTCACCCCTCCACCATGC
AAGATCTCAACACCATGTTGTCGAACACATGACAGCCATTGAAGCCTGTGCCTTCTGGCCC
GGGCTTTGGGCCGGGATGCAGGAGGCAGGCCCCGACCCGTCTTCAGCAGGCCCCCACCTC
CTGAGTGGCAATAAATAAAATCGGTATGCTG

FIGURE 66

MGSGLPLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSYESSFLELLEKLCLLHLPSGTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCGGC
CAGGTGCCCGTCGCAAGGTGCCCTGGCCGGAGATGCGGTAGGAGGGCGAGCGCGAGAAGCCCC
TTCCTCGCGCTGCCAACCCGCCACCCAGCCATGGCGAACCCGGCTGGGCTGCTTCTGGCG
CTGGGCCTGCCGTTCTGCTGGCCGCTGGGCGAGCCTGGGGCAAATACAGACCACCTCTGC
AAATGAGAATAGCACTGTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTTTGGCTGCCTGCTCCTGGCTGTGGGCTG
GCACTGTTGGTGCGGAAGCTTCGGGAGAACGGCAGACGGAGGGCACCTACCGGCCAGTAGCGA
GGAGCAGTTCTCCCATGCAGCCGAGGCCGGCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCCATTAGGTCCCTCTCCTGCATCTGTCTCCCTCATTGCTGTGACCTTGGGAAA
GGCAGTGCCCTCTGGGCAAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAAGAAGGTACTCAA
AGACTCTGCCCTGAGGTCAAGAGAGGATGGGCTATTCACTTTATATTTATATAAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLLALGLPFLARWGRAWGQIQTTSANENSTVLPSSSTSSSDGNLRPEAITAIIVVFS
LLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

FIGURE 70

MGLFRGFVFLVLCLLHQSNNTSFIKLNNNGFEDIVIVIDPSVPEDEKIIEQIEDMVTTASTYLFE
ATEKRFKKNVSILIPENWKENPQYKRPKHENKHADIVVAPPTLPGRDEPYTKQFTECGEKG
YIHFTPDLLLGKKQNEYGPPGKLFVHEWAHLRGVFDEYNEDQPFYRAKSKKIEATRC
SAGISGRN
RVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDKVQTEKASIMFMQSIDS
VVEFCNEKTHNQEAP
SLQNIKCNFRSTWEVISNSEDFKNTIPMVT
PPPPP
VFSLKISQRIVCLVLDKSGSMGGKDRLNR
MNQAAKHFL
LQTVENG
SWVG
MVHF
DSTATIV
NKLIQIKSS
DERNT
IMAGLPTYPLGGTS
ICSGIK
YAFQVIGELHSQLDGSEV
L
LTD
GED
NT
ASS
CIDE
VKQSG
AIVH
FIAL
GRAA
DEAV
IEMS
KITGG
SHFYV
SDEA
QNNGL
IDA
FG
ALT
SG
NT
DLS
QK
SL
Q
LES
KG
LT
LN
SN
AW
MND
TVI
ID
STV
GKD
FFL
IT
WNS
LPPS
ISL
W
DPS
GT
IMEN
FT
V
DAT
SK
MAY
L
S
I
P
G
T
A
K
V
G
T
W
A
Y
N
L
Q
A
K
A
N
P
E
T
L
T
V
T
S
R
A
A
N
S
V
P
P
I
T
V
N
A
K
M
N
K
D
V
N
S
F
P
S
P
M
I
V
Y
A
E
I
L
Q
G
Y
V
P
V
L
G
A
N
V
T
A
F
I
E
S
Q
N
G
H
T
E
V
L
E
L
L
D
N
G
A
G
D
S
F
K
N
D
G
V
Y
S
R
Y
F
T
A
Y
T
E
N
G
R
Y
S
L
K
V
R
A
H
G
G
A
N
T
A
R
L
K
L
R
P
L
N
R
A
A
Y
I
P
G
W
V
V
N
G
E
I
E
A
N
P
P
R
E
I
D
E
T
Q
T
T
L
E
D
F
S
R
T
A
S
G
G
A
F
V
V
S
Q
V
P
S
L
P
D
Q
Y
P
P
S
Q
I
T
D
L
D
A
T
V
H
E
D
K
I
I
L
T
W
T
A
P
G
D
N
F
D
V
G
K
V
Q
R
Y
I
I
R
I
S
A
S
I
L
D
L
R
D
S
F
D
D
A
L
Q
V
N
T
T
D
L
S
P
K
E
A
N
K
E
S
F
A
F
K
P
E
N
I
S
E
E
N
A
T
H
I
F
I
A
I
K
S
I
D
K
S
N
L
T
S
K
V
S
N
I
A
Q
V
T
L
F
I
P
Q
A
N
P
D
D
I
D
P
T
P
T
P
T
P
D
K
S
H
N
S
G
V
N
I
S
T
L
V
L
S
V
I
G
S
V
I
V
N
F
I
L
S
T
T
I

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592, 628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGGAAACCTGGGAGTAGAGTACTGACAGCAAAGACCGGAAAGACCATACGTCCCCGGCAGGGGTGA
CAACAGGTGTATCTTTGATCTCGTGTGGCTGCCTTCTTCAAGGAAAGACGCCAAGGTAATTGACCCA
GAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTCTTGAACCCCAAGTTATGCCAGGATTACTAGAGAGTGTCA
ACTCAACCAGCAAGCGGCTCTCGCTTAACCTTGTTGGAGGAGAACCTTGTGGGCTGCCTCTTAGCA
GTGCTCAGAAAGTGACTTGCCTGAGGGTGGAGAGAACAAAGGAAAGGTCCTCTTGCTGTTGGCTGCACATCAGGAA
GGCTGTGATGGGAATGAAGGTGAAACATGGAGATTCACTTCAGTCATTGCTTGCCTGCAAGATCATCCTTAA
AGTAGAGAAGCTGCTCTGTGGTGTAACTCCAAGAGGAGAACCTGTTCTAGAAGGAAATGATGCAAGCAAGCTC
CGGGGCCCAAACGCATGCTCTGTGGCTAGCCAGGGAAAGCCCTCCGTGGGGCCCGCTTGAGGGATGCC
ACCGGTTCTGGACGCATGGCTGATTCTGAATGATGATGGTCCGCGGGGCTGCTGCGTGGATTCCCGGGTGTG
GTTTGTGCTGGTCTCTCTGTGCTATCTGCTGTACATGTTGGCTGCACCCAAAAGGTGACGAGGAGCAG
CTGGCAGTGGCCAGGCCAACAGCCCCACGGGAAGGGTACAGGGCTCCTCAGGAGTGGAGGAGCAGCAC
CGCAACTACGTGAGCAGCTGAAGGGCAGATCGCAAGCTCAAGGAGAGCTGAGGAGAGGAGTGACGAGCTCAGG
AATGGCAGTACCAAGGCAGCAGTGTGGCTGGTCTGGACAGGAGCCCCAGAGAAACCCAGGCCGACCTC
CTGGCCTCCCTGCACTCGCAGGTGGACAAGGAGAGGTGAATGCTGGCTCAAGTGGCCACAGAGTATGCAGCAGTG
CCTTTGATGATGTTACTACAGAAGGTGACAGCTGGACTGGTACCCCGGAGGAGAACGCTGTG
AGGAAGGACAAGGGATGAGTTGGAGGACATTGAATCAGGCTGGAGACCTGAAACAATCTGCAAGAAACAGC
CCCAATCACCCTTAACGGCTCTGATTCTAGAAGGAGATCAGGAAACAGAAAGGGACAAGGACATTGAT
GAGCTACCTCAAAGGGACCAAAACAGAATTCAAACGGCTCATCTTATTGACCATTCAGCCCCATCATGAAA
GTGAAAAATGAAAGCTCAACATGGCAACACGTTATCAATGTTATGTCCTAGCAAAAGGGTGGACAAGTTC
CGGCACTCATGAGAAATTCAAGGGAGATGTGCAATTGAGCAGGATGGGAGACTGGCCTACATCTACTGTTGTTACTTGGG
AAAGAAGAAATAATGAAGTCAAAGGAATACTGAAACACTTCCAAGCTGCCAACCTCAGGAACCTTACCTTCATC
CAGCTGAATGGAGAATTCTCGGGAAAGGGACTTGTGTTGGAGGACTCTCCACCTCTGGCATGAGAACGGCTCATGGAG
TTTTCTGTGATGTGGACATCTACTTCACATCTGAATTCTCAATACGTTAGGGCTGAATACACAGGAGGGAGAAG
GTATTTATCCAGTCTTCAGTCAGTACAATCTGGCATATAACGGCCACCATGATGCAGTCCCTCCCTTGAA
CAGCAGCTGGTATAAAGAAGGAAATGGAGTTGGAGAGACTGGGATGACGTTGCACTGTTGAG
TTCATCAATATAGTGGTTGATCTGGACATCAAAAGGCTGGGGAGGGATGTGCACCTTATCGCAAGTATCTC
CACAGCAACCTCATGTTGAGGACTCTCCACCTCTGGCATGAGAACGGCTCATGGAG
CTGACCCCGACAGTACAAGATGTCATGCACTGCAAGGCTGAACAGGAGCATCCACGGCCAGCTGGCATGCTG
GTGTTCAAGGACAGAGATAGAGGCTCACCTCGCAAACAGAAACAGAAGACAAGTAGACAAAAAAACATGAACTCCAGA
GAAGGATGTGGAGGACACTTTCTCTCTGGCTGCAATTACTGAAAGTGGCTGCAACAGAGAAAAGACTTCCATAAA
GGACGACAAAAGAATTGGACTGATGGGTCAGAGATGAGAAAGGCTGGGATCTCTGTTGGCTTTTACACAGA
AATCAAATCTCCCTGGCTGCAAAGTAACCCAGTGTGACCTTGTGAAGTGTCTGACAAAGGCAAGATGCTTGT
AGATTATAAGCTAATGGTGTGGAGTTGATGGTTTACAATACACTGAGACCTGTTGTTGTGCTCATTGA
AATATTGATTTAAGAGCAGTTTGTAAAAAAATCTTAGCATGAAAGGCAAGCATATTCTCTCATATGAATGA
GCCATCAGCAGGGCTACTTCTAGGAATCTGAAATGCTTATGAAACACTGGCTTATTGACT
AGTGAGTACATTAAGTAAATAATGGACCAAAAGGAAACCATATAATCTGTCATATTCCCAGAT
TAACCAAAATAATCTGCTTATCTTTGGCTCTTAACTGTCCTCTTCTTCTTAAATGCACT
TTTTCTCTGTGAGTTAGTCCTTATTAATTACACTTGCAGGCCCTACAAGAGGACAAGTTGGCTAC
ATTTTATATTTTAAGAAGATACTTGAGATGCAATTGAGAATCTCAGTCAGGATCAAAGCATCAAATTGATGCCATAT
CCAAGGACATGCCAAATGCTGATTCTGTCAGGCACCTGAATGTCAGGACATTGAGACATAGGGAGGAAATGTTGACT
AATACAGACGTACAGATACTTCTGTAAGGAGTATTTCGAAGGAGTACCTGAGGAGGAAACTGAAACACTGGAGGAAAAGAAAATGAC
ACTTTCTCTTCAAGAAAAGGAAACTCATTGAGACTGTTGATACTGTCATGTCACCTAAAGTCAGAACCCACATT
CTCCTCAGAAGTAGGGACCGCTTCTACCTGTTAAATAACCAAGTATACCGTGTGAACCAAACATCTCTT
AAAACAGGGTGCTCCCTGGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAATATATATATATATATATTGT
GAAAGATCAATCCATCTGCCAGAATCTAGTGGGATGGAAGTTTGTCTACATGTTACCCACCCAGGCCAGGGAG
TAACTGAATTATTTTAAATAATGAGCTTACTCAATCAGGAGTACCTGTTGCAAGTAAACTCATGTTAAAGCTTCAAGAAC
CAAACATTTTAAATAATGAGCTTACTCAATGAGCTTACTCAATCAGGAGTACCTGTTGCAAGTAAACTCATGTTAAAGCTTCAAGAAC
ATTCAAGCTGGTGTGTTAAATGCACTGTTGATGTTGACTGGTAGTTATGAAATTAAATGATTGAG
CCATGAATGGAAGGTGGTATTGCACTAATAATGATTGAGGATATGAA

FIGURE 72

MMMVRRGLLAWISRVVVLLVLLCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEWE
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVRKDKRDELVEAIESALETLNNA
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKEFKRLILFRPFSPIMKVKNEKLMAN
TLININVPLAKRVDKFRQFMQNPREMCIEQDGRVHLTVVYFGKEEINEVKGILENTSKAANFRNF
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFYPVLFQSQY
NPGIIYGHDAVPPLEQQLVIKKETGFWRDFGFGMTCOYRSDFINIGGFDLDDIKGWGGEDVHLYR
KYLHSNLIIVVRTPVRGFLFHLWHEKRCMDELTPEQYKCMQSKAMNEASHGQLGMLVFRHEIEAHL
RKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 73

GAGACTGCAGAGGGAGATAAAGAGAGAGGGCAAAGAGGCAGCAAGAGATTGTCTGGGATCCA
GAAACCCATGATAACCTACTGAACACCGAATCCCTGGAAGGCCACAGAGACAGAGACAGCAAGA
GAAGCAGAGATAAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCACTCCTC
CCTCCCTCTCTGCCTGTCTAGTCCTAGTCCTCAAATTCCAGTCCCAGTCCCAGTCCC
CTGGGACACTATGTTGTTCTCGCCCTCTGCTGGAGGTGATTTGGATCCTGGCTGCAGATGGG
GTCAACACTGGACGTATGAGGGCCACATGGTCAGGACCATTGCCAGCCTTACCTGAGTGT
GGAAACAATGCCAGTCGCCATCGATATTCAAGACAGACAGTGTGACATTGACCCGATTTGCC
TGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCGAGCCTTGGACCTGCACAACAATGCC
ACACAGTGCAACTCTCTGCCCTCTACCCGTATCTGGTGGACTTCCCGAAAATATGTAGCT
GCCAGCTCACCTGCACTGGGTCAAAAGGATCCCCAGGGGGTCAGAACACCAAGATCAACAG
TGAAGGCCACATTGCGAGGCTCACATTGTACATTGACTCTGATTCCTATGACAGCTTGAGTG
AGGCTGCTGAGAGGCCTCAGGGCTGGCTGTCTGGCATCCTAATTGAGGTGGTGAGACTAAG
AATATAGCTTATGAACACATTGAGTCATTGCATGAAGTCAGGCATAAAGATCAGAACACCTC
AGTGCCTCCCTCAACCTAACAGAGAGCTGCTCCCCAACAGCTGGCAGTACTCCGCTACAATG
GCTCGCTCACAACTCCCCCTGCTACCAGAGTGTGCTCTGGACAGTTTTATAGAAGGTCCCAG
ATTCAATGGAACAGCTGGAAAAGCTTCAGGGACATTGTTCTCACAGAAAGAGGAGCCCTCTAA
GCTTCTGGTACAGAACTACCGAGCCCTCAGCCTCTCAATCAGCGCATGGCTTGTCTTCA
TCCAAGCAGGATCCTCGTATACCACAGGTGAAATGCTGAGTCTAGGTGAGGAATCTGGTTGGC
TGTCTGCCTCTCCTGGCTTTATTCATTGCTAGAAAGATTGGAAGAAGAGGCTGGAAAA
CCGAAAGAGTGTGGATGACTCCCTTCACCTCAGCACAGCCACGACTGAGGCATAAATTCCCTCTCAGATAC
CATGGATGTGGATGACTCCCTTCACCTCAGCACAGCCACGACTGAGGCATAAATTCCCTCTCAGATAC
CCAGAAACACTGTAGGAGTAGTAAGCAGATGCTCCTCCCTGACATCTCTTAGAGAGGAAT
GGACCCAGGCTGTCATTCCAGGAAGAAACTGCAGAGCCTCAGCCTCTCAAACATGTAGGAGGAA
ATGAGGAAATCGCTGTGTTAATGCAGAGANCAAACCTGTTAGTTGCAGGGGAAGTTGGG
ATATAACCCCAAAGTCCTCTACCCCTCACTTTATGCCCTTCCCTAGATATACTGCGGGATCT
CTCCTTAGGATAAAGAGTTGCTGTGAAGTTGTATATTGATCAATATATTGAAATTAAAG
TTTCTGACTTT

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPASYPECGNNAQSPIDIQTDSVTFPDLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQINSEAT
FAELHIVHYDSDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLENRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGCTCTGGCGGCCCTGGGACGGCAGTCCCTGTGTC
TCTGGTGGTTGCCAACCTGCAAACATCACCTTATCCATCAACATGAAGAATGCCTACA
ATGGACTCCACCAGAGGGCTTCAAGGAGTAAAGTTACTTACACTGTGCAGTATTCATCACAA
ATTGGCCCACCAGAGGGTGGCACTGACTACAGATGAGAAGTCCATTCTGTTGCTGACAGCTCC
AGAGAAAGTGGAAAGAAAATCCAGAAGACCTCCTGTTCCATGCAACAAATATACTCCAATCTGA
AGTATAACGTGTCTGTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCAC
ACGCTGGTGTCAACCTGGCTGGAGCCGAACACTCTTACTGCGTACACGTGGAGTCCTCGTCCC
AGGGCCCCCTCGCCGTGCTCAGCCTCTGAGAACAGCAGTGTGCCAGGACTTGAAAGATCAATCAT
CAGAGTTCAAGGCTAAATCATCTCTGGTATGTTGCCATATCTATTACCGTGTTCCTTTT
TCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGAAAGAGAAACACCCAGCAAATT
GATTTGATTTATGAAATGAATTGACAAAAGATTCTTGTGCCGTGCTGAAAAAAATCGTGATTA
ACTTTATCACCCCTCAATATCTGGATGATTCTAAAATTCTCATCAGGATATGAGTTACTGGGA
AAAAGCACTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGA
GGAAGAGGAGGGTAAAACATTAGGGTATGCTCGCATTGATGGAAATTGGTGA
AAAACACGGAAGGTACTTCTCACCAGCAAGAGTCCCTCAGCAGAACAAACCCCCGGATAAA
ACAGTCATTGAATATGAATATGATGTCAGAACCAACTGACATTGCGGGGCCCTGAAGAGCAGGA
GCTCAGTTGCAGGAGGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAG
TCTTGGGCCGAAACGTTACAGTACTCATACACCCCTCAGCTCCAAGACTTAGACCCCCCTGGCG
CAGGAGCACACAGACTCGGAGGAGGGCCGGAGGAAGAGCCATCGACGACCCCTGGCGACTGGGA
TCCCCAAACTGGCAGGCTGTGATTCCCTCGCTGTCCAGCTCGACCAGGATTCA
AGCCTCTGAGGGGATGGGCTCGGAGAGGAGGGTCTCTATCTAGACTCTATGAGGAGCCGGCT
CCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCA
TGGAGGAATGGGGTTATA
TGTGCAGATGGAAAACTGATGCCAACACTCCTTGCCTTGTGCAAACAAAGTGAG
TCACCCCTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTCCAGTTGTCA
CTGTGAGAATTACTTATTCTCTATTCTCATAGCAGTGTGATTGGTTCATGCATGTA
GGTCTCTAACAAATGATGGTGGCCTCTGGAGTCCAGGGCTGGCCGGTTCTATGCAGAGAA
AGCAGTCAATAATGTTGCCAGACTGGGTGAGAATTATTCA
GGTGGGTGT

FIGURE 76

MSYNGLHQRFKELKLLTLCISISSQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY
SNLKYNVSVLNTKSNRTWSQCVTNHTLVLWLEPNTLYCVHVESFVPGPPRRAQPSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNEDFKRFFVPAEK
IVINFITLNISDDSKISHQDMSSLGKSSDVSSLNDPQPSGNLRPPQEEEVVKHLGYASHILMEIFC
DSEENTEGTSLTQQESLSRTIPPDKTVIEYEYDVRTTDICAGPEEQELSLOEEVSTQGTLLESQA
ALAVLGPQTLQYSYTPQLQDLDPLAQEHDTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSFDQDS
EGCEPSEGDGLEEGLLSRLYEEPAPDRPPGENETYLMQFMEEWGLYVQMAN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCGAGGACTCCAGCGTCCCCAGGTCTGGCATCCTGCAC TTGCTGCCCTCTGACAC
CTGGGAAG**ATGGCCGGCCCGTGGACCTTCACCCCTCTGTGGTTCTGGCAGCCACCTGATC**
CAAGCCACCCCTCAGTCCCCTGCAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC
ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCATGC
GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGCAGCCTGGTGAACACCGTCTGAAGCACATC
ATCTGGCTGAAGGTCACTCACAGCTAACATCCTCAGCTGCAGGTGAAGCCCTGGCCAATGACCA
GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA
TCGTGGAGTTCCACATGACGACTGAGGCCAAGCCACCATCCGCATGGACACCAAGTGCAAGTGGC
CCCACCCGCTGGCCTCAGTGA CTCAGTGTGCCACCAGCCATGGGAGCCTGCGCATCCAACTGCTGTA
TAAGCTCTCCTCCTGGTGAACGCCCTAGCTAACGAGGTCAAGCAGGTCAACCTCCTAGTGCCATCCCTGC
CCAATCTAGTGA AAAACCCAGCTGTGTCCCGTATCGAGGCTTCAATGGCATGTATGCAGAC
CTCCTGCAGCTGGTGAAGGTGCCATTCCCTCAGCATTGACCGTCTGGAGTTGACCTCTGTA
TCCTGCCATCAAGGGTACACCAATTCAAGCTCACCTGGGGCCAAGTTGGACTCACAGGGAA
AGGTGACCAAGTGGTTCAATAACTCTGCAGCTTCCCTGACAATGCCACCCTGGACAACATCCCG
TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAGCTGCAGTGCTGTGCTCTCCAGAAGA
ATTCAATGGCCTGGACTCTGTGCTTCCCTGAGAGTGCCCCATGGCTGAAGTCAGCATCGGGC
TGATCAATGAAAAGGCTGCAGATAAGCTGGATCTACCCAGATCGTAAGATCTTAACCTCAGGAC
ACTCCCCAGTTTTATAGACCAAGGCCATGCCAAGGTGCCCAACTGATCGTGCTGGAAAGTGT
TCCCTCCAGTGAAGGCCCTCCGCCCTTGTTCACCTGGCATCGAACGCTCGGAAGCTCAGT
TTTACACCAAAGGTGACCAACTTAACTCAACTTGAATAACATCAGCTCTGATGGATCCAGCTG
ATGAACCTGGGATTGGCTGGTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA
CTCCATCCTGCTGCCGAACCAGAAATGGCAAATTAAAGATCTGGGTCCCAGTGTCAATTGGTGAAGG
CCTTGGGATTGAGGCAGCTGAGTCCTCACTGACCAAGGATGCCCTGTGCTTACTCCAGCCTCC
TTGTGGAAACCCAGCTCCTGTCTCCAGT**GAAGACTTGGATGGCAGCCATCAGGAAGGCTGG**
GTCAGCTGGGAGTATGGGTGTGAGCTATAGACCATCCCTCTGCAATCAATAAACACTTG
CCTGTGAAAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVKEKLTQELKDHNATSILQQPLLLSAMREK
PAGGI PVLGSLVNTVLKHI IWLKVITANILQLQVKPSANDQELLVKI PLDMVAGFNTPLVKTIVE
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYLGAKLLDSQGKVT
KWFNNSAASLTMP TLNDNIPFSLIVSQDVVKAAVAAVLSPEEFMVLLDSVLPESAHLRKSSIGLIN
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEALRPLFTLGIEASSEAQFYT
KGDQLILNLNNISSLDRIQLMNSGIGWFQPDVLKNIITEIHSILLPNQNGKLRSGVPVSLVKALG
FEAAESSLTKDALVLTPASLWKPSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCTGAAGACGCTCTACTGAGAGGTCTGCCATGGCCTCT
CTTGGCCTCCAACTGTGGGCTACATCTAGGCCTCTGGGCTTTGGCACACTGGTGCCT
GCTGCTCCCAGCTGAAAACAAGTCTTATGTCGGTGCAGCATTGTACAGCAGTTGGCTCT
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATACCCAGTGTACATCTATAGC
ACCCCTCTGGCCTGCCGCTGACATCCAGGCTGCCAGGCCATGATGGTACATCCAGTGCAT
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGCATGAGATGCACAGTCTCTGCCAGGAATCCC
GAGCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTCATCCTGGAGGCCTCTGGATT
ATTCCTGTTGCCCTGGAATCTCATGGGATCCTACGGGACTCTACTCACCACGGTGCCTGACAG
CATGAAATTGAGATTGGAGAGGCTTTACTTGGGATTATTCCTCCCTGTTCTCCCTGATAG
CTGGAATCATCCTCTGCTTTCTGCTCATCCAGAGAAATCGCTCCAACACTACGATGCTAC
CAAGCCAACCTCTGCCACAAGGAGCTCCAAGGCCTGGTCAACCTCCAAAGTCAAGAGTGA
GTTCAATTCTACAGCCTGACAGGGTATGTGTGAAGAACCCAGGGGCCAGAGCTGGGGGTGGCTG
GGTCTGTAAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACCAACTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTGGCATTGGATTGAGCAAAGGCAGAAATGGG
GCTAGTGTAAACAGCATGCAGGTGAATTGCCAAGGATGCTGCCATGCCAGCCTCTGTTCC
TCACCTTGCTGCCCTGCCCTAAGTCCCCAACCTCAACTGAAACCCATTCCCTTAAGCCA
GGACTCAGAGGATCCCTTGGCTGGTTACCTGGACTCCATCCCCAACCCACTAATCACA
TCCCACTGACTGCCCTGTGATCAAAGACCCCTCTCTGGCTGAGGTTGGCTTAGCTCATT
GCTGGGGATGGAGAGAAGCAGTGGCTTGCTGGCATTGCTCTAACCTACTTCTCAAGCTC
CCTCCAAAGAAACTGATTGCCCTGGAACCTCCATCCACTTTGTTATGACTCCACAGTGTCCA
GACTAATTGTGATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGACATTAAAAAAATA

FIGURE 8o

MASLGLQLVGYILGLLGLLGTILVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMVTTSSAISSLACIIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAWNLLHGLRDFYSPLVPSMKFEIGEALYLGISSLFSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKSEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCACACGCGTCCGCGCCTCTCCCTCTGCTGGACCTTCCTCGTCTCCATCTCCCTCCTTC
CCCGCGTCTCTTCCACCTTCTCTTCCACCTTAGACCTCCCTCCTGCCCTCCTTCCT
GCCCACCGCTGCTTCCGGCCCTCTCCGACCCGCTCTAGCAGCAGACCTCTGGGGTCTGTGG
GTTGATCTGTGGCCCTGTGCCCTCGTGCCTTCTGCTCTCCCTCCTCCGACTCCGCTCCGG
ACCAGCGGCCCTGACCCCTGGGGAAAGGATGGTTCCCGAGGTGAGGGTCCCTCCTCCCTGCTGGGA
CTCGCGCTGCTCTGGTTCCCCCTGGACTCCCACGCTCGAGCCGCCAGACATGTTCTGCCCTT
CCATGGGAAGAGATACTCCCCCGCGAGAGCTGGCACCCCTACTTGGAGGCCACAAGGCCTGATGT
ACTGCCCTGCGCTGTACCTGCTCAGAGGGGCCCATGTGAGTTACCGCCTCCACTGTCCGCCT
GTCCACTGCCCTCAGCCTGTGACGGGCCACAGCAATGCTGTCCCAAGTGTGTGGAACCTCACAC
TCCCTCTGGACTCCGGGCCACCAAAGTCTGCCAGCACACGGGACCATGTACCAACACGGAG
AGATCTCAGTCCCCATGAGCTGTTCCCTCCCGCCTGCCAACCGACTGTGTCCCTCTGCAGCTGC
ACAGAGGGCCAGATCTACTGCCCTCACAAACCTGCCCGAACCAAGGCTGCCAGCACCCCTCCC
ACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGCAAGTGAAGCAATCGGATGAAGAGGACA
GTGTGCAGTCGCTCCATGGGTGAGACATCCTCAGGATCCATGTTCCAGTGTGCTGGAGAAAG
AGAGGCCCGGGCACCCAGCCCCACTGGCTCAGGCCCTCTGAGCTTACCCCTGCCACTT
CAGACCCAAGGGAGCAGGCCAGCACACTGTCAAGATCGTCTGAAGGGAGAAACATAAGAAAGCCT
GTGTGCATGGCGGAAGACGTACTCCCACGGGGAGGTGTGGCACCCGCCCTCCGTGCCCTCGC
CCCTGCCCTGCATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGGCTGGGAAGTGTGCAAGATTGCCAGAGG
ACAAAGCAGACCCGGCCACAGTGAAGATCAGTTTACCAAGGTGTCCCAAGGCACCGGGCCGGTC
CTCGTCCACACATCGGTATCCCAAGGCCAGACAACCTGCCGTGCTTGCCTTGGAACACGAGGC
CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAAACTGAGGCTCAGAGAG
GTGAAGTACCTGGCCAAGGCCACACAGCCAGAATCTCCACTTGACTCAGATCAAGAAAGTCAG
GAAGCAAGACTTCCAGAAAGAGGCACAGCACTCCGACTGCTCGTGGCCCCACGAAGGTCACT
GGAACGTCTTCCAGGCCAGACCCCTGGAGCTGAAGGTACGCCAGTCCAGACAAAGTGACCAAG
ACATAACAAAGACCTAACAGTTGCAGATATGAGCTGTATAATTGTTATTATATTAATAAAA
TAAGAAGTTGCATTACCCCTCAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKCVEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELP
PSRLPNQCVLCSCTEGQIYCGLTTCEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPGTAPPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKHKACVHGGKTY
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTEYPCRHPEKVAGKCKKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLSDQESQEARNPERGTLPTARWPPRRSLERLPSPDGAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTGCCCTCGCTCACCGAGGCCTCTCC
GTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCTCTCTCTAATCCATCCGTACACCTCTCCTGTCA
TCCGTTCCATGCCGTGAGGTCCATTACAGAACACATCCTGGCTCTCATGCTCAGTTGGTTCTGAGTC
TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTGGGCCAGACAAGCCTGTCCAGGCCCTGGTGGGGAG
GACGCAGCATTCTCTGTGTTCTCTAAGACCAATGCAGAGGCCATGGAAGTGGGTTCTTCAGGGG
CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGAAAGGACCAGCCATTATGCAGATGCCACAGTATC
AAGGCAGGACAAAATGGTGAAGGATTCTATTGCGGAGGGCGCATCTCTGAGGCTGGAAAACATTACT
GTGTTGGATGCTGGCTCTATGGTGCAGGATTAGTTCCCAGTCTTACTACCAGAACGGCATCTGGGAGCT
ACAGGTGTCAGCACTGGGCTCAGTCCCTCTCATTTCCATCACGGGATATGTTGATAGAGACATCCAGCTAC
TCTGTCAGTCCTCGGGCTGGTCCCCGGCCACAGCGAAGTGGAAAGGTCACAAGGACAGGATTGTCC
ACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTGATGTGGAGATCTCTGACCGTCCAAGAGAA
CGCCGGGAGCATATCCTGTTCCATGCGGATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
GAGATACCTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGAAATACTCTGCTGTGCCCTA
TTTTTGGCATTGTTGACTGAAGATTTCTCTCCAAATTCCAGTGGAAATCCAGGCCAACTGGACTG
GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCGAAACACGCACTGGGAGGTGACTCTGGATCCAG
AGACGGCTACCCGAAGCTCTGCTTCTGATCTGAAACTGTAACCCATAGAAAGCTCCCCAGGGAGGTG
CCTCACTCTGAGAAGAGATTACAAGGAAGAGTGTGGTGGCTCTCAGAGTTCCAAGCAGGGAAACATTA
CTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGCTGGGAGTGTGGCGGATGATGTGGACAGGAGGA
AGGAGTACGTGACTTTGTCTCCGATCATGGTACTGGTCTCTCAGACTGAATGGAGAACATTGTATTTC
ACATTAATCCCCGTTTATCAGCGCTTCTCCAGGACCCACCTACAAAATAGGGTCTTCTGGACTA
TGAGTGTGGGACCATCTCCTCTCAACATAATGACCAGTCCCTTATTTATACCCCTGACATGTCGGTTG
AAGGCTTATTGAGGCCCTACATTGAGTATCCGCTCTATAATGAGCAAAATGGAACTCCCATAGTCATCTGC
CCAGTCACCCAGGAATCAGAGAACAGGCCCTTGGCAAGGGCCTCTGCAATCCCAGAGACAACCAACAG
TGAGTCCTCCTCACAGGCAACCACGCCCTCCCTCCCCAGGGTGAATGTAGGATGAATCACATCCACAT
TCTTCTTAGGGATATAAGGTCTCTCTCCAGATCCAAAGTCCCGCAGCAGGCCGCAAGGTGGCTTCCA
GATGAAGGGGGACTGGCCTGTCCACATGGGAGTCAGGTGTCATGGCTGCCCTGAGCTGGAGGGAAAGAAGG
CTGACATTACATTAGTTGCTCTCACTCCATGGCTAAGTGATCTTGAATACCACCTCAGGTGAAG
AACCGTCAGGAATTCCCATCTCACAGGTGTGGTGTAGATTAAGTAGACAAGGAATGTGAATAATGCTTAG
ATCTTATTGATGACAGAGTGTATCCTAATGGTTGTTCAATTATATTACACTTCAGTAAAAAAA

FIGURE 84

MALMLSLVSLLLKGSGQWQVFGPDKPVQALVGEDAASFCLSPKTNAEAMEVRFTRGQFSSVH
LYRDGKDQPFMOMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYYQKAIWELQ
VSALGSVPPLISITGYVDRDIQLLCQSSGWFPRPTAKWKGPGQDLSRTNRDMHGLFDVEISL
TVQENAGSISCSMRRAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAELDWRRKHGQAEELRDARKHAVEVTLDPETAHPKLCVSDILKTVTHRKAPQEVPHSEKRF
TRKSVVASQSFQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLSPDHGYWVRLNGEHLYFT
LNPRFISVFPRTPTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT
PIVICPVTQESEKEASWQRASAIPETSNSESSSQATTPFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGGGCCCTGGCACCTCTAACCCAGACATGCTGCTGCTGCTGCTGCCCT
GCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAGAACAGTAACTGCTGACGATGCAGAGATCCGTGA
CGGTGCAGGAAGGCCTGTGTCCATGTGCCCTGCTCCCTCCTACCCCTCGCATGGCTGGATT
TACCCCTGGCCCCAGTAGTCATGGCTACTGGTCCGGAAAGGGCAATACAGACCAGGATGCTCC
AGTGGCCACAAACAACCAGCTGGGAGTGGAGAGACTGGGACCATTCCACCTCCTG
GGGACCCACATACCAAGAATTGCACCCCTGAGCATCAGAGATGCCAGAAGAAGTGATGCGGGGAGA
TACTTCTTCGTATGGGAGAAAGGAAGTATAAAATGGATTATAAAACATCACCGGCTCTGTGAA
TGTGACAGCCTTGACCCACAGGCCAACATCCTCATCCCAGGCACCCCTGGAGTCCGGCTGCC
AGAATCTGACCTGCTGTGCCCTGGCCTGTGAGCAGGGGACACCCCTATGATCTCCTGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCCCTCCTCGGTGCTCACCCCTATCCCACA
GCCCCAGGACCATGGCACCAGCCTCACCTGTCAGGTGACCTTCCCTGGGCCAGCGTGACCACGA
ACAAGACCGTCCATCTCAACGTGCCTACCCGCCTCAGAAACTTGACCATGACTGTCTTCAAGGA
GACGGCACAGTATCCACAGTCTGGAAATGGCTATCTGTCACTCCAGAGGGCCAGTCT
GCGCCTGGTCTGTGCAGTTGATGCAGTGACAGCAATCCCCTGCCAGGCTGAGCCTGAGCTGG
GAGGCTGACCCCTGTGCCCTCACAGCCCTAAACCCGGGGTGCTGGAGCTGCCTTGGGTGCAC
CTGAGGGATGCAGCTGAATTCACCTGCAGAGCTCAGAACCCCTCAGGCTCTCAGCAGGCTACCT
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTCGGGGGAGCTGG
CCACAGCCCTGGCTCTGCCCTGCGTCATCTTCGTTGAGTGAGGCTGCAGGAAGAAA
TCGGCAAGGCCAGCGCAGCGGCGTGGAGATACGGGCATAGAGGATGCAAACGCTGTCAGGGTT
AGCCTCTCAGGGCCCTGACTGAACCTTGGCAGAGACAGCTCCCCCAGACCAGCCTCCCCCAG
CTTCTGCCCGCTCCTCAGTGGGGAGGGAGAGCTCCAGTTATGCATCCCTCAGCTTCCAGATGGT
AAGCCTTGGACTCGGGGGACAGGAGGCCACTGACACCGGAGTACTCGGAGATCAAGATCCAC
AGTGAGAAACTGCAGAGACTCACCCGTATTGAGGGATCACAGCCCCTCCAGGCAAGGGGAGAAGTCA
GAGGCTGATTCTTGTAGAATTACAGCCCTCAACGTGATGAGCTATGATAACACTATGATTATG
TGCAGAGTGAAAAGCACACAGGCTTGTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCC
TCCCTTTATTTTAACTAAAGACAGACAAATTCCTA

FIGURE 86

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTQEGLCVHPCFSYPSHGWIYPGPVHGYWFREG
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLISRARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCSPWACEQGTPPMISWIGTSVSPLDPSTTRS
SVLTLIPQPQDHGTSLTCQVTPGASVTTNKTVHLNVSYPPQNLMTVFQGDGTVSTVLGNSSL
SLPEGQSLRLVCAVDADSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAAEFTCRAQNP
LGSQQVYLNVSLQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGDTGIE
DANAVRGSASQGPLTEPWAEDSPPDQPPPASARSSVGEGELOQYASLSFQMVKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGTGAAGGAGCTCTGTAC
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAAACCAACTCAGCTTCCTGCTGTTTC
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTCAAGGAATGGACCTGT
TCTTCGTCTCCATCTGCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCCTAGTGCATTGA
TGGCCTGTATTTCTCCGCACTGAGAATGGTGTATCTACCAAGCCTCTGTGACATGACCTCTG
GGGGTGGCGGCTGGACCCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGAAAGTGCACGGTG
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGACGGCAACTGGGC
CAACTACAACACCTTGGATCTGCAGAGGCCACGAGCGATGACTACAAGAACCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGCATCTGGCACGTGCCAATAAGTCCCCATGCAGCACTGG
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTCAGACACTGGACATAATCT
GTTTGGCATCTACCAAGAAATATCCAGTGAATATGGAGAAGGAAAGTGTGACTGACAACGGCC
CGGTGATCCCTGTGGTCTATGATTTGGCACGCCAGAAAACAGCATCTTATTACTCACCTAT
GGCCAGCGGAATTCACTGCGGGATTGTTCAAGGGTATTTAATAACGAGAGAGCAGCAA
CGCCTTGTGTGCTGGAATGAGGGTACCGCAGCGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG
GATACTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTCTGGTTGATTGGAGTGGATATCG
GGAACTCATGTTGGTTACAGCAGCCGTGAGATAACTGAGGCAGCAGTGTGCTTCTATTCTATCG
TTGAGAGTTTGTGGAGGGAACCCAGACCTCTCCCAACCATGAGATCCAAAGGATGGAGAA
CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLLFLIAATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGRWSSQQGSKADYPEGDGNWANYNTFGSAEAAT
SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTNDNPVIPHVVYDFGDAQKTASYYSPYGQREFTAGFVQFRVNNEAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTGTCGGCTTGC~~GGGG~~AGACTTCAGGAGTCGCTGTCTGAACCTCCAGCCTCAGAGAC
CGCCGCCCTTGTCCCCGAGGGCCATGGGCGGGCTCAGGGCTTGTGCCCTCTCGCTTCTGACG
CTCCTGGCGCATCTGGTGGTCGTACACCTTATTCTGGTCCC~~GGG~~ACAGCAACATACAGGCC~~T~~
CCTGCCCTCAGTTCACCCCCGAGGAGTATGACAAGCAGGACATT~~CAG~~CTGGTGGCCGCGCTCT
CTGTCACCC~~T~~GGGCCTTTG~~C~~AGTGGAGCTGGCCGGTTCC~~T~~TCAGGAGTCTCCATGTTCAAC
AGCACCCAGAGCCTCATCTCCATTGGGCTACTGTAGTGCATCCGTGGCC~~T~~GCCTTCTTCAT
ATTCGAGCGTTGGGAGTGC~~ACT~~ACGTATTGGTACATTTTG~~T~~CTTCTGCAGTGCC~~TT~~CCAGCTG
TCACTGAAATGGCTTATT~~CGT~~CACCGT~~TT~~GGGCTGAAAAAGAAACCC~~TT~~CTGATTAC~~CTT~~CA
TGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCG~~T~~CGTATTCC~~T~~GAAGAAGGAAG
GCATAGGCTTCGGTTTCCCCTCGAAACTG~~CT~~CTGCTGGAGGATATGTGTTGAATAATTACG
TCTTGAGTCTGGGATTATCCGATTGTATTAGTGC~~TT~~GTAATAAAATGTTTG~~T~~AGTAACA
TTAAGACTTATACAGTTAGGGACAATTAAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFC SALPAVTEMALFV
TVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTGAC
TCGCTGCTGCTTCGTTCTGGTGCAGGGTAGCCTCTATCTGGTATCTGTGCCAGGATGATG
GTCCTCCCGGCTCAGAGGACCTGAGCGTATGACCACGAGGGCAGCCCCGGCCGGTGCCT
CGGAAGCGGGGCCACATCTCACCTAAGTCCGCCCATGGCAATTCCACTCTCCTAGGGCTGCT
GGCCCCGCCTGGGAGGCTTGGGCATTCTGGCAGCCCCCAACCGCCGAACCACAGCCCC
CACCCCTCAGCCAAGGTGAAGAAAATCTTGGCTGGGCACCTACTCCAACATCAAGACGGTG
GCCCTGAACCTGCTCGTCACAGGGAAAGATTGTGGACCATGGCAATGGACCTCAGCGTCCACTT
CCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTCCCCCAGTAAAGCTGTAG
AGTTCCACCCAGGAACAGCAGATCTCATCGAAGCCAAGGCCCTCCAAAATCTTCAACTGCCGGATG
GAGTGGGAGAAGGTAGAACGGGCCCGGACCTCGCTTGACCCACGACCCAGCCAAGATCTG
CTCCCGAGACCACGCTCAGAGCTCAGGCCACCTGGAGCTGCTCCAGCCCTCAAAGTCGTCTGTG
TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAAC
CATAGTGTACCCCCACTACCCATCTGGGTGACCCGGGGCAGGCCACAGAGGCCAGGCCAGGGC
TGGAAAGGACAGGCCCTGCCCATGCAGGAGACCATCTGGACACCGGGCAGGGCAAGGGTTGGGCTC
AGGCAGGGAGGGGGGGTGGAGACGAGGAGATGCCAAGTGGGCCAGGGCAAGTCTCAAGTGGCAG
AGAAAAGGGTCCCAAGTGTGGTCCCAACCTGAAGCTGTGGAGTGA
ACTAGATCACAGGAGCACTGG
AGGAGGGAGTGGGCTCTCTGTGCAGCCTCACAGGGCTTGCCACGGAGCCACAGAGAGATGCTGG
TCCCCGAGGCCCTGTGGGCAGGCCATCAGTGTGGCCCCAGATCAAGTCATGGAGGAAGCTAAGC
CCTTGGTTCTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGAGATT
CATCAGTGTGGACA
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCTTAGGAGCCAGTCAGCAGGGTGGGGTGGG
GCCAGAGGAGCTCCAGCCCTGCCTAGTGGGCCCTGAGCCCCCTGCTGTGCTGAGCATGG
CATGAGGCTGAAGTGGCAACCTGGGCTTTGATGTTGACAGATTGACCATCTGTCTCCAGC
CAGGCCACCCCTTCCAAAATCCCTCTTGTCCAGTACTCCCCCTGTACCACCCATTGCTGATG
GCACACCCATCCTAAAGCTAACAGACAGGAGCATTGTTGCTCCACACTAACGGCCACAGCCCATC
CGCGTGTGTGTGTCCTCTCCACCCCAACCCCTGCTGGCTCCTCTGGAGCATCCATGTCCCG
GAGAGGGTCCCTCAACAGTCAGCCTCACCTGTCA
GAGACCCGGGTTCTCCGGATCTGGATGGC
CGCCCTCTCAGCAGCGGGCACGGGTGGGCCGGCCAGAGCATGTGCTGGATCTGTC
TGTGTGTCTGTCTGTGGGTGGGGAGGGAGGGAAAGTCTGTGAAACCGCTGATTGCTGACTTT
TGTGTGAAGAATCGTGTCTGGAGCAGGAAATAAGCTTGGGGGGCA

FIGURE 92

MQLTRCCFVFLVQGSILYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVANLLVTGKIVDHGNGTF
SVHFQHNATGQGNISISLVPPSKAVEFHQQQIFIEAKASKIFNCRMEEWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPFKVVCVYIAFYSTDYRLVQKCPDNYHSDTPYPPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGCGGCCGTGTTCTCGGCTGCGCCTTCATTGCCTTCGGGCTGCGCTGCC
CTTTATGTCTCACCATGCCATCGAGCCGTTGCGTATCATCTTCATGCCGGAGCTTCTT
CTGGTTGGTGTCTACTGATTCGTCCCTGTTGGTCATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTGAGCGTTGTCTGTCTATATCCAAGAA
ATGTTCCGATTTGATATTATAAACTCTTAAAAAAAGCCAGTGAAGGTTGAAGAGTATAAACCC
AGGTGAGACAGCACCCCTCTATGCGACTGCTGGCCTATGTTCTGGCTGGGCTTGGAAATCATGA
GTGGAGTATTTCCCTTGTGAATACCCCTATCTGACTCCTGGGCCAGGCACAGTGGCATTCAT
GGAGATTCTCCTCAATTCTCCTTATTCAGCTTCATGACGCTGGTCATTATCTGCTGCATGT
ATTCTGGGGCATTGTATTTTGATGGCTGTGAGAAGAAAAAGTGGGCATCCTCCTTATCGTTC
TCCTGACCCACCTGCTGGTGTCAGCCCAGACCTTCATAAGTTCTTATTATGGAAATAACCTGGCG
TCAGCATTATAATCCTGGTGTCAGGGCACCTGGCATTCTAGCTGCGGGGAGGCAGCTGCCG
AAGCCTGAAACTCTGCCTGCTGCAAGACAAGAACACTTCTTACAACCAGCGCTCCAGAT
AACCTCAGGGACCAGCACTCCAAACCGCAGACTACATCTTAGAGGAAGCACAGACTGTGCCT
TTTCTGAAATCCCTTTCTGGTGGATTGAGAAAGAAATAAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAGPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISLWVWMARVIIDNKDG
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV
FSFVNTLSDSLGPCTVGIHGDSQFFLYSAFMTLVIIILHVFWGIVFFDGCEKKWGILLIVLLT
HLLVSAQTFISSYYGINLASAFIILVLMGTWAAGGSCRSILKLCLLCQDKNFLLYNQSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTCACCAGAGTAAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTGTCTTAGGATCAAC
TCGGTCATTACCAACAGCTAAACCTGTTGGACTCCCTCCACAAAACTGGCTCCGGATCAGG
GAACACTACCAAAACCAACAGCAGTCAAATCAGGTCTTCCTTAAGTCTGATACCATTAAACA
CAGATGCTCACACTGGGCCAGATCTGCATCTGTTAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCCCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTTTGTACACAAACTTGGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
TTCACGAGCCTCATCCATTCTGTTCCCGGGAGGCATCCTGCCACCAGTCAGGCAGGGC
TAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAGGAGCAGGTGTAATCCTGCCACCCAGG
GAACCCCAGCAGGCCCTCCAACTCCAGTGGCACAGATGACGACTTGCAGTGACCACCCCT
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCA
GTAAAGCTGTTCAAATTTTCAACTAAGCTGCCTCGAATTGGTGATACATGTGAATCTTATC
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAAATTCTTAATT
TACCTGAAATATTCTGAAATTTCAGAAAATATGTTCTATGTAGAGAATCCAACTTTAAAAA
CAATAATTCAATGGATAAAATCTGTCTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT
TAAAACATATTGGAAAACTGGAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSSLIPLTQM
LTLGPDHLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE
LPQIIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTPAGIQRSTHAIIEEEATTESANGIQ

Signal peptide:

amino acids 1-16

FIGURE 97

GCTCAAGTGCCTGCCTGCCCCACCCAGCCCAGCTGGCCAGAGCCCCCTGGAGAAGGAGCTCTTCTTGCCTGGCAGCTGGACCAAGGGAGCCAGTCTTGGCGCTGGAGGGCTGTCTGACCATG
GTCCTGCCTGGCTGTGGCTGCTTGTGTCCTCCGCTCCCCCAGGCTCTCCCAAGGCCCAGCCTGCAGAGCTGTGAGTCCAGAAACTATGGTGGAAATTCCCTTATACCTGACCAAGTTGC
CGCTGCCCGTGAGGGGCTGAAGGCCAGATCGTGTGCTCAGGGACTCAGGCAAGGCAACTGAGGGCCATTGCTATGGATCCAGATTCTGGCTCTGCTGGTACCCAGGGCCCTGGACCGAGAGGA
GCAGGCAGAGTACCACTACAGGTACCCCTGGAGATGCAAGGATGGACATGTCTTGTGGGGTCCACAGCCTGTCACGTGAGTGGACATCTCAAGGCAACTTAC
AGCCTGTGCTGTGACGTGAAGGATGAGAAATGACCAAGGTGCCCCATTCTCTCAAGGCAACTTAC
AGAGCTGGCTGAGCCGGGTAACAGGCCATGGCATTCCCTCTTGTGAGGCTCAGACCGGGATGAGCCAGGCTCCAGCCAGC
GGATGAGCCAGGCACAGCCAACCTGGATCTTCACATCCAGGCCACATGGCCAGGTA
CTTCCCCAGACATGTTCCAGCTGGACCTCGGCTGGGGCTCTGGCCCTCAGCCCCAAGGGAGC
ACCAGCCTGACCAAGGCCATGGAGAGGACCTACAGCTGTTGGTACAGGTCAAGGACATGGGTGA
CCAGGCCTCAGGCCACAGGCCACTGCCACCGTGGAAAGTCTCCATCATAGAGAGCACCTGGGT
CCCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGTCTATACCCGACCATGGCCAGGTA
CACTGGAGTGGGGGTATGTGCACTATCACCTGGAGAGCCATCCCCGGGACCTTTGAAGTGAA
TGCAGAGGGAAACCTCTACGTGACCAAGAGACTGGACAGAGAACGCCAGGCTGAGTACCTGCTCC
AGGTGGGGCTCAGAATTCCCATGGCGAGGACTATGGGCCCCCTCTGGAGCTGCACGTGCTGGTG
ATGGATGAGAAATGACAACGTGCTATCTGCCCTCCCGTACCCACAGTCAGCATCCCTGAGCT
CAGTCCACCAAGGTACTGAAGTACTAGACTGTCAGCAGAGGATGCAAGATGCCCGGCTCCCCA
ATTCCACGTTGTATCAGCTCTGAGCCCTGAGCCTGAGGATGGGTAGAGGGAGAGCCTTC
CAGGTGGACCCACTTCAGGCAGTGTGACGCTGGGGCTCCACTCGAGCAGGCCAGAACAT
CCTGCTCTGGTGTGGCATGGACCTGGCAGGGCAGAGGGTGGCTCAGCAGCAGTGTGAAG
TCGAAGTCGCACTCACAGATATCAATGATCACGCCCTGAGTTCACTTCCAGATTGGCCT
ATAAGCCTCCCTGAGGATGTGGAGCCGGACTCTGGTGGCATGCTAACAGCCTATTGATGCTGA
CCTCGAGCCCGCCTCCGCCTCATGGATTGCAATTGAGAGGAGACAGAAGGGACTTTG
GCCTGGATTGGGAGCCAGACTCTGGCATGTTAGACTGACTCTGCAAGAACCTCAGTTATGAG
GCAGCTTCAAGTCATGGGTGGTGGTGGTGCAGAGTGTGGCAAGACTGGTGGGGCCAGGCC
AGGCCCTGGAGCCACGCCACGGTACTGCTAGCTTCCCATCAGTCCCCAGCCGCTCTTCTGCTGACCATC
ACCAGGAGAGCTACGAGGCCAGTGTCCCATCAGTCCCCAGGGTCAATGACTCAGAGGGCTGGCT
CAGGCCCTGGGAGGTGACACCGCCAGTCCCTGCAAGGGGCCAGCCTGGGG
CTGCAATTGAGAAATTCTCGGGGAGGTGACACCGCCAGTCCCTGCAAGGGGCCAGCCTGGGG
ACACCTACACGGTGTGTTGGAGGCCAGGATACAGCCCTGACTCTTGGCCCTGTGCCCTCCAA
TACCTGCAACACCCGCCAACGACCATGGCTGATGTCAGTGGACCCAGCAAGGACCCGATCT
GGCCAGTGGCAGGGTCCCTACAGCTTACCCCTGGTCCCAACCCACGGTGCACAGGGATGGC
GCCCTCAGACTCTCAATGGTCCCATGCCCTACCTCACCTGGCCCTGATTGGTGGAGGCCAGT
GAACACATAATCCCGTGGTGTGAGCCACAATGCCAGATGTGGCAGCTCTGGTTCAGTGAT
CGTGTGCTGCAACGTTGGAGGGGCAGTGTGCACTGGCAGGTGGGGCGCATGAAGGGCATGCCCA
CGAAGCTGTCGGCAGTGGCATCTTGTAGGCACCCCTGGTAGCAATAGGAATCTTCTCATCTC
ATTTTACCAACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTGCCCT
GAAGGCAGACTGTCAGTGGCCAGGCAGCTAGCTGGAGCTTGGCCTCTGGCTCCATCTGAG
TCCCTGGGAGAGAGGCCAGCACCAAGATCCAGCAGGGACAGGACAGAGTAGAAGGCCCTCCA
TCTGCCCTGGGTGGAGGCACCATCACCACGAGGATGTCTGCAAGGCCAGGACACCAACTT
TATGGACTGCCATGGAGTGTCCAAATGTCAAGGGTGTGTTGCCAATAATAAGCCCCAGAGAA
CTGGGCTGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 98

MVPAWLWLLCVSVPQALPKAQPAELSVEVPENYGGNFPYLTKLPLPREGAEGQIVLSGDGKAT
EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHFSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQPSPDMFQLEPRLGALALSPKG
STSLDHALERTYQLLVQVKDMGDQASGHQATATVEVSIIESTWVSLPEPIHLAENLKVLYPHHMAQ
VHWSSGGDVHYHLESHPPGPFEVNAEGNLYVTRELDREAAQEYLLQVRAQNSHGEDYAAPPLEHVL
VMDENDNVPICPPRDPTVSIPELSPPGTEVRLSAEDADAPGSPNSHVYQLLSPEPEDGVEGRA
FQVDPTSGSVTLGVPLRLAGQNILLVLAMDLAGAEGGFSSTCEVEVAVTDINDHAPEFITSQIG
PISLPEDVEPGTLVAMLTAIDADLEPAFRIMDFAIERGDTEGTFGDWEPDSGHVRLRLCKNLSY
EAAPSHEVVVVVQSVAKLVGPGPGPAGATATVTVLVERVMPPPQLDQESYEASVPISAPAGSFLLT
IQPSDPISRTLRFSLVNDSEGWLICIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALTLPVPS
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRDWRLQTLNGSHAYLTLALHWEP
REHIIIPVVVSHNAQMWQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGTLVAIGIFLI
LIFTHWTMSRKKDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGGCTAAGGAACCCAGGCATCCAGCTGCCACGCC
AGTCCAAGATTCTTCCCAGGAACACAAACAGTAGGAGACCCAGCCTCTGGAAGCACCAGCCTTA
TCTCTCACCTCAAGTCCCCTTCTCAAGAATCTCTGTTCTTGCCTCTAAAGTCTTGGTAC
ATCTAGGACCCAGGCATCTTGCTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAATG
TTCTCCTTATGTTGGTCTACTATTGCATTAGAGCTGAACAAATTCCATGAGACTAGCACC
TCTGCCAACACTGGATCCAGTGTGATCTCAGTGGAGCCAGCACAGCCACCAACTCTGGTCCAG
TGTGACTCCAGTGGGTCAAGCACAGCCACCATCTCAGGGTCCAGCGTACCTCCAATGGGTCA
GCATAGTCACCAACTCTGAGTCCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
TTCAGCACAGCGTCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACACCTCCAGTGG
GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCCTCAGTGGGCCAGCACAGTCACCAACT
CTGGTCCAGTGTGACCTCCAGTGGAGCCAGCAGTGCACCAACTCTGAGTCCAGCACAGTGTCC
AGTAGGGCCAGCAGTGCACCAACTCTGAGTCTAGCAGCACACTCTCCAGTGGGCCAGCACAGCC
CAACTCTGACTCCAGCACAACCTCCAGTGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGCCACCAACTCTGAGTCCAGTGG
GCCACCAACTCTGAGTCCAGCACACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAG
AACGACCTCAATGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACAGCCACCTCCAGTGGGCC
GCACAGCCACCAACTCTGACTCCAGCAGCACAGCCACCAACTCTGAGTCCAGCACAGCCACCAACTCTGAG
TCCAGCACGACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGCCACCTCCAGTGG
GGCTAGCACAGCCACCAACTCTGAGTCCAGCACACCTCCAGTGGGCCAGCACAGCCACCAACT
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACACCCCTCC
AGTGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGCCAACACAGCCAC
CAACTCTGAGTCCAGCAGTCAGTGTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGTCAGCACAGCCACCAACTCTGAGTCCAGCACACCTCCAGTGGGCCAGCAC
GCCACCAACTCTGACTCCAGCACAACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCTAG
CACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACACCTCCAGTGGGCC
ACACAGCCACCAACTCTGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGA
ATGCACCAACTCTCCATAGTCATCTACTGAGTCCAGTGAGTGGCAAAGCTGGTGGGTCCCTGGT
GCCGTGGGAAATCTCCTCATCACCCTGGTCTCGGTTGTGGCGCCGTGGGCCCTTGTCTGGC
TCTTCTCTGTGAGAACAGCAGCTGGCTTGGAGGAAACACCTTTAACACAGCTGTCTACCACCC
CATGGCTCAACCATGGCCTGGCTGGAGGAGACCAACTCTGAGTGGAGGAAATCATGGAGCCCC
GTGGAGTCCTAAGTGGTCTGGAGGAGACCAACTCTGAGTGGAGGAAAGAGACCTGGCA
ACAGCAGCCCCTGAGCAGCCCCGGAGCAAGTGCAGCATTCTCAGGAAGGAAGAGACCTGGCA
CCCAAGACCTGGTTCCCTTCATTCTCACCTTCTGCTTAAACACAGCTGTCTGGAA
AATCTGAAGAAGGTATTCTCACCTTCTGCTTAAACACAGACACTGGAAGAGAAACTATAT
TGCTCATTTAGCTAAGAAATAACATCTCATCTAACACACAGCACAAAGAGAAAGCTGTGCTTG
CCCCGGGGTGGGTATCTAGCTCTGAGATGAACCTAGTTATAGGAGAAAACCTCCATGCTGGACTC
CATCTGGCATTCAAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKM**Q**GNVLLMF**Q**LLLH**Q**EAATNSNETSTSANTGSSVISSGASTATNGSSVTSSGVSTATISGS
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSG**Q**ISIATNSESSTTSSGASTATNSESSTPSS
GASTVTNGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNDSSTTSSGASTA
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTSNGAGTATNSES
STTSSGASTATNDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNDSSTTSS
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSTTSSGVSTATNSESSTTSSGASTATNDSSTTSSEASTATNSESSTVSSGISTVTNSES
STTSSGANTATNGSSVTSAGSGTAALTGMHTTS**Q**ASTAVSEAKPGGSLVPWEIFLITLVVVA
AVGLFAGLFFCVRNLSLRNTFNTAVYH**Q**HGLNHGLGP*Q*PGGNHGAPHRPRWSPNWFWR*Q*PVSSI
AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCCTCCCGTTACGGGATGAATTAACGGCGGGTCCGCACGGAGGTTGTGACCCCTA
CGGAGCCCCAGCTGCCACGCCACACTCGCGTCGCCGCGTGCCTGCTGTACAGGTG
GGAGGCTGGAACATATCAGGCTGAAAAACAGAGTGGGTACTCTCTGGAAAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGAAAGGGAAATTGTGGTGCTCTGAACCCATGGTCAATT
AACGAGGCAGTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAGCTTGGAAATCAT
GGTGTCAGGAAAGGGATTTACTTATACTGACTCTGTTGGGAAGCTTTGGAAAGCATT
TCATGCTGAGTCCCTTTTACCTTGATGTTGTAACCCATCTGGTATCGCTGGATCAACAAAC
CGCCTTGTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTGGTGTAAAGT
GATTATAACTGGGATGCATTGTTCTGGAGAAAGAAGTGTCAATTATCATGAACCACCGAACAA
GAATGGACTGGATGTTCTGTGGAATTGCTGATGCCATAGCTACCTCAGATTGGAGAAATT
TGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTGGTGGCCATGCAGGCTGCTGCCTATAT
CTTCATTCAAGGAAATGGAAGGATGACAAGAGCATTGCAAGACATGATTGATTACTTTGTG
ATATTCAACGAACCACTCAACTCCTCATATTCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTGCTGAAAAAAATGGACTTCAGAAATATGAATATGTTTACATCCAAG
AACTACAGGCTTACTTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACCTTGATGCTGTCCATG
ATATCACTGTGGCGTATCCTCACAAACATTCCCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTT
CCCAGGGAAATCCACTTCACGTCCACCGGTATCCAATAGACACCCCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGTGCACAAACGGTGGGAAGAGAAAGAAGAGGAGGCTGCCTCCTATCAAG
GGGAGAAGAATTTTATTTACCGGACAGAGTGTCAATTCCACCTGCAAGTCTGAACCTCAGGGTC
CTTGTGGTCAAATTGCTCTATACGTATTGGACCTGTCAGCCCTGCAATGTGCCTACTCAT
ATATTGTACAGTCTGTTAAGTGGTATTTATAATCACCATTGTAATCTTGCTGCAAGAGA
GAATATTGGTGGACTGGAGATCATAGAACCTGCATGTTACCGACTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAATGAGTAAGATTATAAGGTTGCCATGTGAAAACCTAGAGCATAATTG
GAAATGTTCTAAACCTTCTAACGCTCAGATGCATTGATGACTATGTCGAATATTCTTACT
GCCATCATTATTGTTAAAGATATTTGCACTTAACCTTGTGGAAAAATATTGCTACAATT
TTAATCTCTGAATGTAATTGATACTGTGTACATAGCAGGGAGTGATGGGGTGAATAACTT
GGGCCAGAATATTAAACAAATCATCAGGCTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSFGIMVSWKGIFYFILTLFWGSFFGSIFMLSP
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM
FLWNCLMRYSYLRLEKICLKASLKGVPGFWAMQAAAYIFIHRKWDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSKRSNAFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDITVA
YPHNIPQSEKHLQGDGPREIHFHVHRYPIDTLPLTSKEDLQLWCHKRWEEKERLRSFYQGEKNF
YFTGQSVIPPCCKSELRLVLLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLOERIFGG
LEIIELACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCCACGGCTCTGCGCTGAGACAGCTGGCTGACC
TCCAAATCATCCATCCACCCCTGCTGTCATCTGTTCATAGTGTGAGATCAACCCACAGGAATA
TCCATGGCTTTGTCATTTGGTTCTCAGTTCTACGAGCTGGTGTAGGACAGTGGCAAGT
CACTGGACCGGGCAAGTTGTCAGGGCTTGGTGGGGAGGACGCCGTGTCCTGCTCCCTCT
TTCCTGAGACCAGTGCAGAGGCTATGGAAGTGCAGGTTCTCAGGAATCAGTCCATGCTGTGGTC
CACCTCTACAGAGATGGGGAAAGACTGGGAATCTAACGAGATGCCACAGTATCGAGGGAGAACTGA
GTTTGTGAAGGACTCCATTGCAAGGGGGCGTGTCTCTAAGGCTAAAAAACATCACTCCCTCGG
ACATGGCCTGTATGGGTGTCAGTCCCAGATTACGATGAGGAGGCCACCTGGGAGCTG
CGGGTGGCAGCACTGGGCTCACTTCCCTCATTTCCATCGTGGGATATGTTGACGGAGGTATCCA
GTTACTCTGCCTGTCTCAGGCTGGTTCCCCCAGCCCACAGCCAAGTGGAAAGGTCCACAAGGAC
AGGATTGTCTCAGACTCCAGAGCAAATGCAGATGGGTACAGCCTGTATGATGTTGGAGATCTCC
ATTATAGTCCAGGAAAATGCTGGGAGCATATTGTTCCATCCACCTGTCAGCAGAGTCATGA
GGTGGAAATCCAAGGTATTGATAGGAGAGACGTTTCCAGCCCTCACCTGGCCCTGGCTCTA
TTTACTCGGGTTACTCTGTTGCCCCCTGTTGTTGATGATAATTGTTCTTC
AAATCCAAGGAAATCAGGGGAACTGGACTGGAGAAGAAAGCACGGACAGGCAGAATTGAG
AGACGCCCGAACACGCACTGGAGGTGACTCTGGATCCAGAGACGGCTCACCGAAGCTCTGCG
TTTCTGATCTGAAGACTGTAACCCATAGAAAAGCTCCAGGAGGTGCTCACTCTGAGAAGAGA
TTTACAAGGAAGAGTGTGTTGCTTCAGGGTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
CGTGGGACAAATGTAGGGTGGTATGTTGGAGTGTGTCGGGATGACGTAGACAGGGGAAAGAAC
ATGTGACTTGTCTCCAAACAATGGGTATTGGTCTCAGACTGACAACAGAACATTGTATTTC
ACATTCAATCCCCATTATCAGCCTCCCCCAGCACCCCTCTACACGAGTAGGGGCTTCT
GGACTATGAGGGTGGGACCATCTCCTCTCAATACAAATGACCAAGTCCCTTATTATACCTGC
TGACATGTCAGTTGAAGGCTTGTGAGACCCCTATATCCAGCATGCCATGTATGACGGAAAAG
GGGACTCCCATATTCAATGTCAGTGTCTGGGATGAGACAGAGAACCCCTGCTTAAAGGGC
CCCACACCAGACCCAGACACAGCCAAGGGAGAGTGTCCCGACAGGTGGCCCCAGCTTCT
CCGGAGCCTGCGCACAGAGACTCACGCCCTTCTAGGGAGCTGAGGGTCTTCTGCC
TGAGCCCCGAGCAGGGCAGTCACAGCTTCCAGATGAGGGGGATTGGCTGACCCGTGGGAG
TCAGAAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCACATTAGGTTAGTTGTGAAA
CTCCATCCAGCTAACAGCATTGAACAAGTCACAACCTCCAGGCTCTCATTTGCTAGTCACGG
ACAGTGATTCCCTGCCCTCACAGGTGAAGATTAAAGAGACAACGAATGTAATCATGCTTCAGGTT
TGAGGGCACAGTGTGCTAATGATGTTTTATATTACATTTCACCCATAAACTCTGTT
TGCTTATTCCACATTAATTACTTTCTCTATACAAATCACCCATGGAATAGTTATTGAACACC
TGCTTGTGAGGCTAAAGAATAAAGAGGGTAGGATTTCACTGATTCTATAAGCCCAGCAT
TACCTGATACCAAAACCAGGCAAGAAAACAGAAGAAGAGGAAGGAAACTACAGGTCCATATCC
CTCATTAACACAGACACAAAAATTCTAAATAAAATTAAACAAATTAAACTAAACATATATTTA
AAGATGATATATAACTACTCAGTGTGGTTGTCCCAAAATGCAGAGTTGGTTAATATTTAAAT
ATCAACCAGTGTAAATTCAAGCACATTAATAAAAGTAAAAAGAAAACCATAAAAAAAAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFIRNQFHAVVH
LYRDGEDWESKQMPQYRGRTEFKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLLCISSLGWFQOPTAKWKGPGQGDLSSDSRANADGSLYDVEISI
IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQSPWRLASILLGLLCGALCGVVMGIIIVFFK
SKGKIQAELDWRRKHGQAEELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAPQEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNVGWWVGCRDDVDRGKNNVTLSNNNGYWVRLLTTEHLYFT
FNPHFISLPPSTPPTRGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLLRPYIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTCATTGCTGGTGGCAATGATGTATCGGCCAGATGTGGTAGGGCTAGGAAAAGAG
TTTGTGGAACCTGGTTATCGCCTCGTCATCTTCATATCCCTGATTGTCTGGCAGTGTGCATTGGA
CTCACTGTTCAATTATGTGAGATATAATCAAAGAACCTACAATTACTATAGCACATTGTCATTTAACAC
TGACAAACTATATGCTGAGTTGGCAGAGAGGCTTAACAATTACAGAAATGAGCCAGAGACTTGAAT
CAATGGTCAAAATGCATTTATAATCTCCATTAAAGGGAAAGAATTGTCAAGTCTCAGGTATCAAGITC
AGTCAACAGAACATGGAGTGTGGCTCATATGCTGTTGATTGTAGATTCACTCTACTGAGGATCCTGA
AACTGTAGATAAAATTGTTCAACTGTTTACATGAAAGCTGCAAGTGTGAGGACCCCTAAAGTAG
ATCCTCACTCAGTTAAATTAAAAAAATCAACAAGAACAGAACAGACAGCTATCTAAACCATGCTGCCA
ACACGAAGAAGTAAAACCTCTAGGTCAAGAGTCTCAGGATCGTGGTGGACAGAACAGTAGAACAGGGTGAATG
GCCCTGGCAGGCTAGCCTGCAGTGGATGGGAGTCATCGTGTGGAGCAACCTTAATTATGCCACATGGC
TTGTGAGTGCTGCTCACTGTTTACAACATATAAGAACCTGCCAGATGGACTGCTTCCTTGGAGTAACA
ATAAAACCTCGAAAATGAAACGGGTCTCCGAGAAATAATTGTCCATGAAAATACAAACACCCATCACA
TGACTATGATATTCTCTTGCAAGAGCTTCTAGCCCTGTCCTACACAAATGCACTAGATAGAGTTGTC
TCCCTGATCCATCCTATGAGTTCAACCAGGTGATGATGTTGTGACAGGATTGGAGCACTGAAAAT
GATGGTTACAGTCAAAATCATCTCGACAAGCACAGGTGACTCTCATAGACGCTACAACATTGCAATGAACC
TCAAGCTTACAATGACGCCATAACTCCTAGAAATGTTATGCTGGCTCCTAGAAGGAAAACAGATGCAT
GCCAGGGTCACTGGAGGACACTGGTTAGTCAGATGCTAGAGATATCTGGTACCTGCTGGAATAGTG
AGCTGGGGAGATGAATGTGCGAACACCAACAAGCCTGGTTTATAGACTAGAGTTACGGCCTGCGGGACTG
GATTACTTCAAAACTGGTATCTAAGAGACAAAAGCCTCATGGAACAGATAACATTTTTTGTTGTTG
GGTGTGGAGGCCATTAGAGATAACAGAAATTGGAGAACACTTGCAAAACAGCTAGATTGACTGATCTCA
ATAAACTGTTGCTTGATGCACTGATTTCTCCAGCTGTTCCGCACGTAAAGCATCCTGCTTCTGCCA
GATCAACTCTGTCATCTGTGAGCAATAGTTGAAACTTTATGTCATAGAGAAATAGATAATACAATATTAC
ATTACAGCCTGATTCAATTGTTCTAGAAGTTGTCAGAATTGACTGACATAATTGTAAT
GCATATATAACAATTGAAGCACTCCTTTCTCAGTTCTCAGCTCCTCTCATTTCAAGCAAATATCCATT
TCAAGGTGCAAAACAGGAGTGAAGAAAATATAAGAAGAAAAAAATCCCTACATTATTGGCACAGAA
AAGTATTAGGTGTTCTTAGGAAATTAGAAATGATCATATTCAATTGAAAGGTCAAGCAAAGACA
GCAGAAATACCAACTCATTTAGGAAGTATGGAAACTAAGTTAAGGAAGTCCAGAAAGAACCAAG
ATATATCCTTATTTCATTTCCAAACAACTACTATGATAAAATGTAAGAAGATTCTGTTTGTGACCT
ATAATAATTACAAACTCATGCAATGTAATTGTTCTAAGCAAATTAAAGCAAATTATTAACATTG
TTACTGAGGATGTCAACATATAACAAATAAAATATAACACCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWIVLVI FISLIVLAVCIGLT VHYVRYNQKKTYNYYSTLSFTTDKLY
AEFGREASNNFTEMSRQLES MVKNAFYKSPLREFVKSQVIKFSQQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTE TDSYLNHCCGTRRSKTLGOSLRIVGG
TEVEEGEWPWQASLQWDGSHRCGATLINATLVSAAHCF TYKNPARWTASFGVTIKPSKMKRL
RRIIVHEKYKHP SHDYDISLAELSSPVPYTN A VRVCLPDASYE FQPGDVMFVTGFGALKNDGYS
QNHLRQAQVTLIDATT CNEPQAYNDAITPRMLCAGSLEGKTDACQGDGGPLVSSDARDI WYLAG
IVSWGDEC A KPNKPGVYTRVTALRDWITSK TGI

Transmembrane domain:
amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCCGAAGAAGTCCCTG
CCCCGATGAGCCCCCGCGTGCCTCCCGACTATCCCAGGCAGGCGTGGGGCACCGGGCCCAGC
GCCGACGATCGCTGCCGTTGCCCTGGGAGTAGGATGTTGAAGGATGGGGCTCTCCCT
ACGGGGCTCACAATGGCCAGAGAAGATTCCGTGAAGTGTCTGCCTGCCTACGCCCTCAA
TCTGCTCTTGGTTAATGTCCATCAGTGTGTTGGCAGTTCTGCTGGATGAGGGACTACCTAA
ATAATGTTCTCACTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTGACTTACTTCC
GTGGTTCATCCGGTCTGATTGCTGTTGCTGTTCCCTATCATTGTTGGGATGTTAGGATATTG
TGGAACGGTGAAAGAAATCTGTTGCTTCTGCATGGTACTTGGAGTTGCTGTCATTCT
GTGAGAACTGGCTGTGGCCTGGACATATGAACAGGAACCTATGGTCCAGTACAATGGTCA
GATATGGTCACTTGAAGGCCAGGATGACAAATTATGGATTACCTAGATATCGGTGGCTACTCA
TGCTTGAATTTCAGAGAGGTTAAGTGTGTTGGAGTAGTATATTCACTGACTGGTGG
AAATGACAGAGATGGACTGGCCCCCAGATTCCTGCTGTTAGAGAATTCCAGGATGTTCCAAA
CAGGCCACCAGGAAGATCTCAGTGCACCTTATCAAGAGGGTTGTGGGAAGAAAATGATTCC
TTTGAGAGGAACCAAACACTGCAGGTGCTGAGGTTCTGGGAATCTCCATTGGGTGACACAAA
TCCTGGCCATGATTCTCACCATACTCTGCTCTGGGCTGTATTATGATAGAACGGAGCCTGG
ACAGACCAATGATGTCCTGAAGAATGACAACCTCAGCACCTGTCTGTCAGTAGAACT
GTTGAAACCAAGCCTGTCAAGAATCTTGACACACATCCATGGAACAGCTTAATACACACT
TTGAGATGGAGGAGTTATAAAAGAAATGTCACAGAAGAAAACCAACAAACTTGTTTATTGGACT
TGTGAATTGGAGTACATACTATGTTCAAGAAATATGAGAAATAAAATCTGCCATAAAA
TAACACCTAACATATACTATTCTATGCTTAAATGAGGATGGAAAAGTTCATGTCATAAGTC
ACCACCTGGACAATAATTGATGCCCTAAATGCTGAAGACAGATGTCATACCCACTGTGTAGCC
TGTGTATGACTTTACTGAACACAGTTATGTTGAGGCAGCATGGTTGATTAGCATTTCCGA
TCCATGCAAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGGTTGATTACTTCAACAA
CTAGTATATAAAAGTACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATTACTTATT
CTCAGCGATCTATTCTCTGATGCTAAATAATTATGCTAACATAGGAATTTCAATATTGGTACT
ACCTAAATGTTGATTTGCTGGTTACTAAATAATTCTTCAACTTAAAGAGCAAGCTAACACAT
TGTCTTAAGCTGATCAGGGATTTTGTATATAAGCTGTGTTAAATCTGTATAATTCACTGAT
TTCAGTTCTGATAATGTTAAGAATAACCTATTGAAAAGGAAAATTGTCCTGTATAGCATCATT
ATTTTAGCTTCTGTTAATAAAAGCTTACTATTCTGCTGGCTTATATTACACATATAAC
TGTTATTAAATACTTAACCACTATTGAAAATTACCACTGTGATACATAGGAATCTTAC
AGAATGTTAGCTGGCTTTAGGAAGTTATAAGAAAATTGTCACATAACTTAGTTGATTCAA
AAGGACTTGTATGCTGTTTCTCCAAATGAAGACTTTGACACTAAACACTTTTAAAAA
GCTTATCTTGCCTCTCCAAACAAAGCAATAGTCTCCAGTCAATATAAAATTCTACAGAAAA
TAGTGTCTTTCTCCAGAAAATGCTGTGAGAATCATTAAACATGTGACAATTAGAGATT
CTTGTGTTATTCACTGATTAATATACTGTCGCAAATTACACAGATTATTAATTTTACAA
GAGTATAGTATATTGAAATGGAAAAGTCATTTACTGTATTTGTGTTAGAGAGTAA
TTCTCAGAATATGAAAGAAAATTAAATGTCATAAAATTTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAWSAWMRDYLNNVLTATRVEEAVILTYFPVVHP
VMIAVCCFLIIVGMLGYCGTVKRNLNLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMVT
LKARMNTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTEMWDPPDSCCVREFPGCSKQAHQ
EDLSDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDRREPQTDQM
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGGACACCTTATCCCACTCATCCTCATCCTTCTCTGATAAAGCCCCTACCAGTGCT
GATAAAGTCTTCTCGTAGAGCCTAGAGCCCTTAAAAAAAAAGTGCCTGAAAGAGAAGGGGACAAAGGAACA
CCAGTATAAGAGGATTTCCAGTGTTCCTGGCAGTGGTCCAGAAGGATGCCTCATTCCCTGCTTCTCACCTG
CCTCTTCATCACAGGCACCTCCGTGTCACCCGTGCCCTAGATCCTGCTGCTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACAGTGGATGAGTCAGGTCTCTATGTGACAACCAGTGAATGGGAG
TGGTACCACTTCAGGGCATGGCGGAGATGCCATGCCACCTCTGCATACAGAAAACCACTGTGGAACCC
CCGACCTGTCTGGCTCAATGCCAGCCACCCCTAGAAGGCAGGGATTGTGCAACGCCAGGCTGTGCCAGCT
TCAATGGGAACGTCTGTCTCTGGAACACACCGTGGAAAGTCAAGGCTTGCCCTGGAGGCTACTATGTGTATCGT
CTGACCAAGCCCAGCGCTGCTTCCACGCTACTGTGGTCAATTGACATCTGCGACGAGGACTGCCATGG
CAGCTGCTCAGATACCAGCGAGTGCACATGCCACCTGAGGACTGTGCTAGGCCCTGACAGGCAGACATGCTTG
ATGAAAATGAATGTGAGCAAACACGGTGGCTGAGTGTGAGATCTGTGAAACCTCAAAACTCCTACCGCTGT
GAGTGTGGGTTGGCGTGTGCTAAGAAGTGTGCAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAA
TGGTGGCTGCAGCCACTTGCCTGGATCTGAGAAAGGCTACCGAGTGTGAATGTCCCCGGGCCTGGTGT
CTGAGGATAACCACACTTGCCAAGTCCCTGTGTTGTGCAAATCAAATGCCATTGAACTGAACATCCCCAGGGAG
CTGGTTGGTGGCCTGGAGCTTCCCTGACCAACACCTCTGCCAGGGAGTGTCCAACGGCACCCATGTCAACAT
CCTCTCTCTCAAGACATGTGGTACAGTGGTCACTGGTGAATGACAAGATTGTGGCAGCAACCTCGTGA
CAGGTCTACCCAAGCAGACCCGGGGAGCAGGGGACTTCATCATCGAACAGCAAGCTGCTGATCCCCGT
ACCTGCGAGTTCCACGCCGTACACCATTCTGAAGGATACGTTCCAACCTTCGAAACTCCCCACTGGAAAT
CATGAGCCGAAATCATGGGATCTCCCATTCACTCTGGAGATCTCAAGGACAATGAGTTGAAGAGCCTTAC
GGGAAGCTCTGCCACCCCTCAAGCTCGTGAACCTCTACTTGGCATTGAGCCCTGGTGCACGTGAGCGGC
TTGGAAAGCTTGGTGGAGAGCTCTTGCCACCCCACTCCAAGATGACGAGGCTCTGAAATACTACCTCAT
CCGGGATGGCTGTGTTCAGATGACTCGGTAAGCAGTACACATCCCAGGATCACCTAGCAAAGCACTCCAGG
TCCCTGTCCTCAAGTTGTGGCAAAGACCACAAGGAAGTGTCTGCACTGCCGGTTCTGTGTGGAGTG
TTGGACAGCGTCTCCCTGTGCCACGGTTGCCACCGGAATGCGTCTGGGCAGGAGGAGACTCAGC
CGGTCTACAGGGCCAGACGCTAACAGCGGCCGATCCGCATCGACTGGAGGACTAGTTGTAGCCATACCTC
GAGTCCCTGCATTGGACGGCTCTGCTCTGGAGCTCTCCCCCACCAGGCTCTAAGAACATCTGCCAACAGC
TGGGTTCAGACTTCACACTGTGAGTCAGACTCCCAGCACCACACTCACTCTGATTCTGGTCCATTCACTGG
CAGGTACAGCACTGCTGAACAATGTGGCCTGGTGGGTTCATTTCTAGGGTTGAAAACATAACTGTCCA
CCCAGAAAGACACTCACCCATTCCCTCATTTCTTCACTTAAACACCTCGTGTATGGTGCAATCAGAC
CACAAAATCAGAAGCTGGTATAATATTCAAGTACAAACCCAGAAAAATAAACAGTTACTGAAATTATGA
CTTAAATACCAATGACTCCTTAAATATGAAATTATAGTTACCTGAAATTCAATTCAAATGCAGACTAA
TTATAGGAATTGGAAGTGTATCAATAAAACAGTATATAATT

FIGURE 110

MPPFLLLTCFITGTSVSPVALDPCSAVISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFTGMAGDAMP
TFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCLWNTTVEVKACPGGYYVYRLTKPSVCFHV
YCGHFYDICDEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSEICVNLKNSYRCECGVGRV
LRSDGKTCEDVEGCHNNNGCShSCLGSEKGYQCECPRLVLSEDNHTCQVPVLCKSNAIEVNIPRELVGG
LELFILTNTSCRGVSNGTHVNILFSLKTCGTVVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFITLEIFKDNEFEEPYREALPTLKLRDLSLYFGIEPVVHV
SGLESLVESCFTATPTSKIDEVLKYLLIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVGKDHKEVFLHCRV
LVCGVLDERSRCAQGCHRRMRRGAGGEDSAGLQQQTLTGGPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGCGTGAGGGACCAAGGCCTGCCCTGCACACTGG
GCCTCCTCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGAGGCCCT
CCTGCTGCCCTGGGGTACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT
TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTGATGTCAAACCCCTGCGCAAACCCGTATCCCC
ATGGAGACCTTCAGAAAGGTGGGATCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTTG
GGTTGCTCTCATCAAGGTGATTCTGGATAAAATACTACTTCCCTGCGGGCAGCCTCTCCACTTCATCCGA
GGAAGCAGCTGTGACGGAGAGCTGGACTGTCCTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC
CCCGAAGGGCCTGCACTGGCAGTCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC
AGGGAACTGGTTCTCGCCTGTTGACAACCTTCACAGAAGCTCGCTGAGACAGCCTGTAGGCAGATGG
GCTACAGCAGAGCTGTGGAGATTGGCCAGACCAAGGATCTGGATGTTGAAATCACAGAAAACAGCCAG
GAGCTTCGCATGCGGAACCTCAAGTGGCCCTGTCCTCAGGCTCCCTGGCTCCCTGCACTGTCTTGCCTG
TGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGGGAGGAGGACATCCTGGACCCCCACTGGTCTCACGGCA
TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGTCTCACGGCA
GCCCACTGCTTCAGGAAACATCCGATGTGTTCACTGGAAAGGTGCGGGCAGGCTCAGACAAACTGGCAG
CTTCCCCATCCCTGGCTGTGGCAAGATCATCATCATTAATTCAACCCCATGTACCCCCAAAGACAATGACA
TCGCCCTCATGAAGCTGCAGTCCCCTCACTTCTCAGGCACAGTCAGGCCATCTGTCTGCCCTTCTT
GATGAGGAGCTCACTCCAGCCACCCACTCTGGATCATGGATGGGCTTACGAAGCAGAATGGAGGGAA
GATGTCAGACATACTGCTGCAGGCGTCAGTCAGGTCATTGACAGCACACGGTGCAATGCAAGCAGATGCGT
ACCAGGGGAAGTCACCGAGAAGATGATGTGTCAGGCACATCCCGGAAGGGGTGTGGACACCTGCCAGGGT
GACAGTGGTGGGCCCCCTGATGTAACCAATCTGACCAAGTGGCATGTGGTGGCATCGTTAGCTGGGCTATGG
CTGGGGGGCCCGAGCACCCAGGAGTATACACCAAGGTCAGCCTATCTCAACTGGATCTACAATGTCT
GGAAGGCTGAGCTGTAATGCTGCTGCCCTTGCAGTGTGGAGGCCCTCCCTGCCCTGCCACCT
GGGATCCCCAAAGTCAGACACAGAGCAAGAGTCCCCTGGTACACCCCTGCCCCACAGCCTCAGCAT
TTCTGGAGCAGCAAAGGCCCTCAATTCTGTAAGAGACCCCTCGCAGCCAGGGCGCCAGAGGAAGTCA
GCAGCCCTAGCTCGGCCACACTTGGTGTCCAGCATCCCAAGGGAGAGACACAGCCCACTGAACAAGGTCT
CAGGGGTATTGCTAAGCCAAGAAGGAACCTTCCCACACTACTGAATGGAAGCAGGCTGTCTGTAAAAGCC
CAGATCACTGTGGCTGGAGAGGAGAAGGAAAGGGTGTGCCAGCCCTGTCCGTCTCACCCATCCCCAA
GCCTACTAGAGCAAGAAACCAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT
ACTGTTGTCATTGTTATTACAGCTATGCCACTATTATAAGAGCTGTGTAACATCTGGAAAAAAA
AAAA

FIGURE 112

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIALLSIASIIIVVLIKVILDKYYFLCG
QPLHFI PRKQLCDGELDCPLGEDEEHCVKS FPEGPAVAVRLSKDRSTLQVLD SATGNWFSACFDN
FTEALAE TACRQMGSRAVE IGPQDLDVVE ITENSQELRMRNSSGPCLSGSLVSLHCLACGKSL
KTPR VVGEEASVDSWPWQVSIQYDKQHVC GGSILD PHWVLTA AHCFRKHTDVF NWKVRAGSDKL
GSFPSLAVAKIIIEFNPMYPKDNDIALMKLQFPLTFSGT VRPICL PFFDEELTPATPLWIIGWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDT CQGD SGGPLMYQS
DQWHVVGIVSWG YGCGGPSTPGVYTKV SAYLNWIYNVWK AEL

Transmembrane domain:
amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCTGGTCCAAGTGATCCACCCGCCTCAGCCTCCAAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTCTGCCCTGAACAACCTTTCACTGCAACTAAAAAGCCACAGGAGT
TGAACGTCTAGGATTCTGACTATGCTGTGGCTAGTGCTCCTACTCCTACCTACATTAAAATC
TGTTTTTGTCTTGTAACTAGCCTTACCTCCTAACACAGAGGATCTGTCAGTGTGGCTCT
GGCCCAAACCTGACCTCACTCTGGAACGAGAACAGAGGTTCTACCCACACCGTCCCCTGAAG
CCGGGGACAGCCTCACCTGCTGGCCTCTCGCTGGAGCAGTGCCTCACCAACTGTCTACGTCT
GGAGGCAGTCACTCGGGCAGTGCAGGTAGCTGAGCCTCTGGTAGCTGCGGCTTCAAGGTGGC
CTTGCCCTGGCGTAGAAGGGATTGACAAGCCGAAGATTCACTAGGCATGGCTCCACTGCC
AGGCATCAGCCTGCTGTAGTCATCACTGCCCTGGGCCAGGACGGGCCGTGGACACCTGCTCA
GAAGCAGTGGGTGAGACATCACGCTGCCGCCATCTAACCTTCACTGCTGCACATCACCTG
ATCCATGGGCTAACATCTGAACCTGTCCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
CAGAAGGGGCTGCTTAGACCACCTGGTTATGTGACAGGACTTGCAATTCTGGAAACATGAGG
GAACGCCGGAGGAAGCAAAGTGGCAGGGAGGAACCTGTGCCAAATTATGGGTCAAGAAAAGATG
GAGGTGTTGGTTATCACAAGGCATCGAGTCTCCTGCATTCACTGAGCTGACAGGACTGTGGCTCAGACC
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGCCATCAGACAGCCGTTCCGCCCGAT
CCACGTACCGCTGCTGAAGGGCAACTGCAGGCCATGCTCTCATCAGCCAGGCAGCAGCCAAA
TCTGCGATCACCAGCCAGGGCAGCCGTCTGGAAAGGAGCAAGCAAAGTGACCAATTCTCCTCC
CTCCTCCCTCTGAGAGGCCCTCTATGTCCCTACTAAAGCCACAGCAAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTGGCCAGGAGGTAGCAAGGCCCTGAGAGCTGATCAGAAGGGCTGCT
GTGCGAACACGGAAATGCCCTCAGTAAGCACAGGCTGCAAATCCCCAGGCAAAGGACTGTGTGG
CTCAATTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGGTT
CAAATGATCTCAAGGCCCTTATACCCAGGAGACTTGATTTGAATTGAAACCCCAAATCCA
AACCTAAGAACCAAGGTGCATTAAGAACATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACAT
TTGGGAGGCCAGGGGGTAGATCACCTGAGGTAGGAGTTCAAGACCAGCCGGCTGGCAACATGG
TGAAACCCCTGTCTACTAAAAAATACAAAAAAACTAGCCAGGCATGGTGGTGTGCCTGTATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTGAAACCTGGAGGTGAAGGAGGCTGAGACA
GGAGAACATTCAGCCTGAGCAACACAGCGAGACTCTGTCAGAAAAAATAAAAAAGAATT
TGGTTATTGTAA

FIGURE 114

MLWWLVLLLPTLKSVFCSLVTSLYLPNTEDSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRD

Signal peptide:

amino acids 1-15

FIGURE 115

CAGCAGTGGTCTCTAGCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCTGGCAA
AGAATCCTCCAGAGAATTGTGAAGACTGTACATTCTAAATGCAGAAGCTTTAAATCCAAGAAA
ATATGTAAATCACTTAAGATTGTGGACTGGTGTGGTATCCTGGCCCTAACTCTAAATTGTCTT
GTTTGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGATTTACATGAAATTGATCCTGTGACCAGAACTGAAATA
TTCAGAAGCGGAAATGGCACTGATGAAACATTGAAAGTGCACGACTTTAAAACGGATACTGG
CATCTACTCGTGGGTCTCAAAATGTTTATCAAAACCTAGATTAAGTGATTCTGAATT
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCAACTTCTTGAAACAGTCAGTG
ATTGGGTCCCAGCAGAAAGCCTATTGAAACCGAGATTCTTCTTAAATTCCAAAATTCTGGA
GATTTGTGATAACGTGACCATGTATTGGATCAATCCCCTCTAATATCAGTTCTGAGTTACAAG
ACTTTGAGGAGGGAGGAGAAGATCTTCACTTCTGCCAACGAAAAAAAGGGATTGAACAAAAT
GAACAGTGGGTGGTCCCTCAAGTGAAGTAGAGAAGACCCGTACGCCAGACAAGCAAGTGAGGA
AGAAACTCCAATAATGACTATACTGAAATAGAATTGATCCCATGCTGGATGAGAGAG
GTTATTGTTGTATTACTGCCGTCAGGCACCCGCTATTGCCGCCGTCTGTGAACCTTTACTA
GGCTACTACCCATATCCACTGCTACCAAGGAGGACGGAGTCATCTGCGTGTATGCCTTG
TAACTGGTGGGTGGCCCGCATGCTGGGGAGGGTCTAATAGGAGGTTGAGCTCAAATGCTTAAAC
TGCTGGCAACATATAATAATGCATGCTATTCAATGAATTCTGCCTATGAGGCATCTGGCCCT
GGTAGGCCAGCTCTCCAGAATTACTGTAGGTAATTCTCTCTTATGTTCTAATAAAACTTCTACA
TTATCACCAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFSKSKICKSLKICGLVFGILALTLLFWGSKHFWPEVPKKAYDME
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLEVHDFKNGYTGIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEIITTFEQSVIWVPAEKPIENRDFLKNKILEICDNVTMYWINPTLISVSE
LQDFEEEGEDLHFPAKEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMED
ERGYCCIYCRRGNRYCRRVCEPLLGYYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCTCAGGAGCGCGTTAGCTTCACACCTCGCAGCAGGGGGCAGCTCTCGCAGGCCA
GGCGGGCGGCCAGGATCATGTCCACCACATGCCAAGTGGTGGCGTCCCTGTCCATCCTGGGCT
GGCGGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCAGGACCTGTACGACAACCCCGTCACCT
CCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTCGTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCC
TATTTACCACATCCTGGACTTCCAGCCATGTCAGGCAGTGCGAGCCCTGATGATCGTAGGCATCGTCT
GGGTGCCATTGGCCTCTGGTATCCATCTTGCCTGAAATGCATCCGATTGGCAGCATGGAGGACTCTG
CCAAAGCCAACATGACACTGACCTCCGGATCATGTTCAATTGTCTCAGGTCTTGTGCAATTGCTGGAGTG
TCTGTGTTGCCAACATGCTGGTACTAATTCTGGATGTCCACAGCTAACATGTACACCGGATGGGTGG
GATGGTGCAGACTGTTCAAGACCAGGTACACATTGGTGCCTGTCGTTGGCTGGTCGCTGGAGGCC
TCACACTAACATTGGGGGTGTGATGATGTGCATGCCCTGCCGGGCTGGCACCAGAAGAAACCAACTACAAA
GCCGTTCTTATCATGCCCTCAGGCCACAGTGTGCTACAAGCCCTGGAGGCTCAAGGCCAGCACTGGCTT
TGGGTCCAACACCAAAAAACAAGAAGATATACGATGGAGGTGCCGCACAGAGGACGAGGTACAATCTATC
CTTCCAAGCAGCACTATGTGTAATGCTCTAACGACCTCTCAGCACGGCGGAAGAAAACCTCCGGAGAGCTCA
CCCCAAAAACAAGGAGATCCCATCTAGATTCTCTGCTTTGACTCACAGCTGAAAGTTAGAAAAGCCT
CGATTTCATCTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTGTCTCAAATATTCCACCATAAAACA
GCTGAGTTATTATGAATTAGAGGCTATAGCTCACATTTCATCCTCTATTCTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGGTTAATCTCTCTCACATTGATGATTAGACAGACTCCCCCTC
TTCCCTCTAGTCATAAAACCCATTGATGATCTATTCCAGCTATCCCCAGAAAACCTTTGAAAGGAAA
GAGTAGACCCAAAGATGTTATTCTGCTGTTGAATTGTCTCCCCACCCCAACTGGCTAGTAATAA
ACACTTACTGAAGAAGCAATAAGAGAAAGATATTGTAATCTCTCCAGCCCATGATCTGGTTTCTT
ACACTGTGATCTAAAGTTACCAACCAAAAGTCATTTCTAGGAGTTCCCTGAGCTCTCCACTGGAGTCCTTTCTT
TTGACATCTCTTATTACAGCAACACCATTCTAGGAGTTCCCTGAGCTCTCCACTGGAGTCCTTTCTT
CGCGGGTCAGAAATTGCTCTAGATGAATGAGAAAATTATTTTAAAGCTAAATATAAGTTAA
AATAAAATAATGTTAGTAAAATGATACACTATCTCTGTGAAATAGCCTACCCCTACATGTGGATAGAAG
GAAATGAAAAAATAATTGCTTTGACATTGTCTATATGGTACTTTGTAAGTCATGCTTAAGTACAATTCC
ATGAAAAGCTCACACCTGTAATCTAGCATTGGGAGGCTGAGGAGGAAGGACTCTGAGCCAGAAGT
TCGAGACTAGCCTGGCAACATGGAGAAGCCCTGTCTCACAAATACAGAGGAAAAATCAGCCAGTC
TGGTGGCATACACCTGTAAGTCCAGCATTCCGGGAGGCTGAGGTGGGAGGACTTGAGCCCAGGGAGGT
TGGGGCTGCAGTGAGGCCATGATCACACCACTGCACTCCAGGAGGTGACATAGCGAGATCCTGTCTAAAAA
AATAAAAATAATAATGAAACACAGCAAGTCCTAGGAAGTAGGTTAAAACATAATTCTTAA

FIGURE 118

MSTTCQVVAFLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLWRSCVRQSSGFTECRP
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGVVAGGLTLIGGVMMCIA
CRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYP SKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAAACTGTTCTCTGTGGCACAGAGAACCTGCTCAAAGCAGAAGTAGCAGTTCCGGAGTCC
AGCTGGCTAAACTCATCCCAGAGGATAATGGCAACCCATGCCTAGAAATCGCTGGCTGTTCTTG
GTGGTGTGGAATGGTGGCACAGTGGCTGTCACTGTCATGCCTCAGTGGAGAGTGTGCGGCCTTCATT
GAAAACAACATCGTGGTTTGAAAACCTCTGGAAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
CATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTTTCTCCGGACCTACAGGCAGCCAG
GACTGATGTGTGCTGCTCCGTGATGTCCTTCTGGCTTCATGATGCCATCCTGGCATGAAATGC
ACCAGGTGCACGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTCAT
CATCACGGGCATGGTGGCTCATCCCTGCTGAGCTGGGTTGCCAATGCCATCATCAGAGATTTCTATA
ACTCAATAGTGAATGTTGCCAAAAACGTGAGCTGGAGAAGCTCTACTTAGGATGGACCACGGCA
CTGGTGCTGATTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTGTTGCAACGAAAAGAGCAGTAGCTA
CAGATACTGATAACCTCCCATCGCACAACCCAAAAAGTTATCACACCGGAAAGAAGTCACCGAGCG
TCTACTCCAGAAGTCAGTATGTTAGTTGTATGTTTTTAACTTAAAGCCATGCAAATG
ACAAAAAATCTATATTACTTTCTCAAAATGGACCCCAAGAAAACTTGATTTACTGTTCTTAACGCCT
AATCTTAATTACAGGAACGTGCATCAGCTATTATGATTCTATAAGCTATTCAGCAGAATGAGATA
TTAAACCAATGCTTGATTGTTCTAGAAAGTATAGTAATTGTTCTAAGGTGGTCAAGCATTCTA
CTCTTTTATCTTACTTCAAAATGACATTGCTAAAGACTGCATTATTTACTACTGTAATTCTCC
ACGACATAGCATTATGTACATAGATGAGTGTAAACATTATCTCACATAGAGACATGCTTATATGGT
TTTATTAAAATGAAATGCCAGTCCATTACACTGAATAAATAGAAACTCAACTATTGCTTTCAGGGAA
ATCATGGATAGGGTGAAGAAGGTTACTATTAAATTGTTAAAAACAGCTTAGGGATTAATGTCTCCA
TTTATAATGAAGATTAAAATGAAGGCTTAATCAGCATTGTAAGGAAATTGAATGGCTTCTGATAT
GCTGTTTTAGCCTAGGAGTTAGAAATCCTAACTCTTATCCTCTTCCAGAGGCTTTTT
CTTGTGTATTAAAATTAAACATTAAAACGCAGATATTGTCAAGGGCTTGCATTCAAAGCTT
TTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAGTGTGTTAGGAAAGTG
AAAATTTTTGTTTGTATTTGAGAAGAATGATGCAAGGAAATCATATATGTATGGAT
ATATTAAATAAGTATTGAGTACAGACTTGGAGTTCATCAATATAAAAGAGCAGAAAAATA
TGTCTGGTTTCTTACCAAAAAACACAACAAAAAAAGTGTCTTGTGAGAACACTCACCT
GCTCCTATGTGGGTACCTGAGTCAGTCAAAATTGTCATTGTTCTGTGAAAATAATTCTCTGTA
CCATTCTGTTAGTTACTAAAATCTGAAATACTGTATTTCTGTTATTCAAATTGATGAA
ACTGACAATCCAATTGAAAGTTGTGTCGACGTCTGCTAGCTTAAATGAATGTGTTCTATTGCTT
TATACATTATTAATAAAATTGTACATTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTAVTVMPQWRVSIFIENNIVVFENFWEGLWMNCVRQANIRMQCK
IYDSLLALSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRCTGDNEKVKAHILLAGIIFIITG
MVVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTALVLIVGGALFCCVFCCNEKSSY
RYSIPSHRTTQKSYHTGKKSPSVYRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGCGGGTGAAGGCAGATTGATGCAGCCTGCGCGGCCTCGGAGCGCGGGAG
CCAGACGCTGACCACGTTCCCTCCTCGGTCTCCTCCGCCCTCCAGCTCCCGCTGCCGGCAGCC
GGGAGCCATGCGACCCCAGGGCCCCGCCCTCCCGCAGCGGCTCCGCGGCCCTCCTGCTGCTCC
TGCTGCTGCAGCTGCCCGCGCGTCAAGCGCCTCTGAGATCCCAAGGGGAAGCAAAGGCCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGCCAATGTTATTCCGGTACACCTGGATCCCAGGTGGGATG
GATTCAAAGGAGAAAAGGGGGATGTCTGAGGGAAAGCTTGAGGGAGTCCTGGACACCCAACCTAC
AAGCAGTGTTCATGGAGTTCAATTGAATTATGGCATAGATCTGGAAAATTGCGGAGTGTACATT
TACAAAGATGCGTCAAATAGTGTCTAAGAGTTGTCAGTGGCTCACTTCGGCTAAATGCA
GAAATGCATGCTGTCAGCGTTGGTATTCAATTCAATGGAGCTGAATGTTCAAGGACCTCTCCC
ATTGAAGCTATAATTATTGGACCAAGGAAGGCCCTGAAATGAATTCAACAATTAAATTATCATCG
CACTTCTCTGTGGAAGGACTTGTGAAGGAATTGGTGGATTAGTGGATGTTGCTATCTGG
TTGGCACTTGTTCAGATTACCCAAAAGGAGATGCTTCACTGGATGGAATTCAAGTGGCTCGCATC
ATTATTGAAGAACTACCAAAATAATGCTTAATTTCATTTGCTACCTCTTTTATTATGCC
TTGGAATGGTTCACTTAAATGACATTAAATAAGTTATGTATAACATCTGAATGAAAAGCAAAG
CTAAATATGTTACAGACCAAAGTGTGATTTCACACTGTTTAAATCTAGCATTATTCAATTG
CTTCAATCAAAAGTGGTTCAATATTTTTAGTTGGTTAGAATACTTCTCATAGTCACATT
CTCTCAACCTATAATTGGAATATTGTTGGTCTTTGTTCTCTTAGTATAGCATTGTTA
AAAAAAATATAAAAGCTACCAATTTGTACAATTGTAATGTTAAGAATTGTTATATCTGT
TAAATAAAAATTATTCCAACA

FIGURE 122

MRPQGPAAASPQRLRGLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAGVPGR
DGSPGANVIPGTPGIPGRDGFKGEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK
MRSNSALRVLFGSLLRKCRNACCQRWYFTFNGAECSGPLPIEAIYYLDQGSPEMNSTINIHRTS
SVEGLCEGIGAGLVDVIAIWVGTCSDYPKGDASTGWNSVSRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGCTCTCTGCCTTCTGGCTCCAACGCAGCTGTGGCTGAA
CTGGGTGCTCATCACGGGAACTGCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
ATTGCCTGGAAGAATACATCATGTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTTTTA
ACCGCCCCCTCCCCACCCCCCAAAAAACTGTAAGAGTCAAAACGTAATATCCATGAAGATCC
TATTACCTAGGAAGATTTGATGTTGCTGCGAATGCGGTGTTGGGATTATTGTTCTGGAG
TGGTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCCAAGGGGCTCAATT
TCTTCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGGCTGTCATGCAACTG
GCCCTAAGCCAAAGCAAAAGACCTAAGGACGACCTTGAACAATACAAAGGATGGGTTCAATG
TAATTAGGCTACTGAGCGGATCAGCTGAGCACTGGTTAGGCTTACGCCCCACTGTTACTGACAATG
CTTTCTCTGCCAACGAGGATGCCCTAAGGGCTGTAGGTGTGAAGGCAAATGGTATATTGTA
ATCTCAGAAATTACAGGAGATACCCCTAAGTATATCTGCTGGTGCCTAGGTTGTCCTCGCT
ATAACAGCCTCAAAACTTAAGTATAATCAATTAAAGGGCTCAACCAGCTCACCTGGCTATAC
CTTGACCATACCATATCAGCAATATTGACAAAATGCTTTAATGGAATACGCAAGACTCAAAGA
GCTGATCTTAGTCTCAATAGAATCTCTATTTCTAACATACCTCAGACCTGTGACAAATT
TACGGAACCTGGATCTGCTCTATAATCAGCTGCATTCTGGATCTGAAACAGTTGGGGCTT
CGGAAGCTGCTGAGTTACATTACGGCTAACCTCCCTGAGAACCATCCCTGTGCGAATATTCA
AGACTGCCAACCTGGAACCTTGGACCTGGGATATAACCGGATCCGAAGTTAGCCAGGAATG
TCTTGTGGCATGACTCAAAGAACCTCACCTGGGACACAATCAATTTCAGCTCAAC
CTGGCCCTTTCAAGGTTGGCAGCCTCAGAACCTTACTTGCAGTGGAAATAAAACTAGTGT
CATAGGACAGACCATGTCCTGGACCTGGAGCTCCTAACAAAGGCTTGATTTATCAGGCAATGAGA
TCGAAGCTTCAGTGGACCCAGTTTCCAGTGTGTCCTCGAATCTGCAGCGCCTCAACCTGGAT
TCCAACAAGCTCACATTATTGGTCAAGAGATTTGGATTCTGGATATCCCTCAATGACATCAG
TCTTGCTGGGAATATATGGGAATGCAAGCAGAAATTTGCTCCCTGTAAACTGGCTGAAAAGTT
TTAAAGGTCTAAGGGAGAATACAATTATCTGTCAGTCCCAAAAGAGCTGCAAGGAGTAAATGTG
ATCGATGCACTGAGAAACTACAGCATCTGTCAGCTTAAAGCCCAAGCTCCCAGGCCGAAGCATGAG
GGCTCTCCAAAGCCGACGTTAACGCCAACAGGCCAGAGACCGATGCTGACGCCAGCACATCT
TGCCCCCGACGGTGGGAGCCACAGAGGCCAGAGACCGATGCTGACGCCAGCACATCT
TTCCATAAAATCATCGCGGGCAGCGTGGCCTTTCTGTCCGTGCTCGTCACTCTGCTGGTTAT
CTACGTGTCTGGAGCGTGGTACCCCTGCGAGCATGAAGCAGCTGCAAGCGCTCCCTCATGCGAA
GGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCAGCACCCAGGAATTATGTA
GATTATAACCCACCAACAGGAGACCGAGATGCTGCTGAATGGGACGGGACCCCTGCACCTA
TAACAAATCGGCTCCAGGGAGTGTGAGGTATGAACATTGTGATAAAAAGAGCTTAAAGCT
GGGAAATAAGGGCTTTATTGAACCTGGTCACTATCAAGGGAAACGCGATGCCCTCCCTCCCC
TTCCCTCTCCCTCTACATTGGTGGCAAGATCCTCCTGTCCGTTTAGTGCATTCAAAACT
GGTCATTTCTCTCATACATAATCAACCCATTGAAATTAAATACCACAATCAATGTGAAGCTT
GAACCTCCGGTTAATATAATACCTATTGTATAAGACCTTACTGATTCCATTAAATGTCGCAATT
GTTTAAGATAAAACTCTTTCATAGGTAAAAAAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTVLLTMLSSAERGCPKGCRCEGKMYCESQKLQEIPSSISAGCLG
LSLRYNLSQKLKYNQFKGLNQLTWLYLDHNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR
PVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHRSNSLRTIPVRIQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKELHLEHNQFSKLNALFPRLVSLQNLQWNKISVIGQTMSTWSSLQRLDL
SGNEIEAFSGPSVFQCVPNLQRLNLDNSNKLFIGQEILDWSWISLNDISLAGNIWECSRNICSLVN
WLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRKHE
SKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVIIVYVSWKRYPASMKQLQQR
SLMRRHRKKRQSLKQMTPSTQEFYVDYKPTNTETSEMLLNNGTGPCTYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTGCCTACTGCTGAATGTCCGTCCCAGGAGGAGGGCTTTGCCGCTG
ACCCAGAGATGGCCCGAGCGAGCAAATTCTACTGTCCGGCTGCGCGCTACCGTGGCGAGCT
AGCAACCTTCCCTGGATCTCACAAAAACTCGACTCCAAATGCAAGGAGAAGCAGCTTGCTC
GGTTGGGAGACGGTCAAGAGAATCTGCCCTATAGGGGAATGGTGCACAGCCCTAGGGATC
ATTGAAGAGGAAGGCTTCTAAAGCTTGGCAAGGAGTGCACACCCGCCATTACAGACACGTAGT
GTATTCTGGAGGTCAATGGTCACATATGAACATCTCCGAGAGGTGTGGCAAAAGTGAAG
ATGAGCATTATCCCCTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTGTATTGGCCAGTT
TTAGCCAATCCAACTGACCTAGTGAAGGTTCAGATGCAAATGGAAGGAAAGGAAACTGGAAGG
AAAACCATTGCGATTCGTGGTGTACATCATGCATTGCAAAATCTTAGCTGAAGGAGGAATAC
GAGGGCTTGGCAGGCTGGTACCCAAATATAACAAAGAGCAGCACTGGTGAATATGGGAGATT
ACCACTTATGATACGACTGAAACACTACTTGGTATTGAATACACCACTTGAGGACAATATCATGAC
TCACGGTTATCAAGTTATGTTCTGGACTGGTAGCTCTATTCTGGAACACCAGCCATGTCA
TCAAAAGCAGAATAATGAATCAACCACGAGATAAACAAAGGAAGGGACTTTGTATAAATCATCG
ACTGACTGCTTGATTCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTACC
ATCTTGGCTGAGAATGACCCCTGGTCAATGGTGTCTGGCTTACTTATGAAAAAATCAGAGAGA
TGAGTGGAGTCAGTCCATTTAA

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMQGEAALARLGDGARES
APYRGMVRTALGIIEEEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLREVVFGKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMQMEGKRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCGLVASILGTPADVIKSRI MNQP
RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCAAGCAGGTGGCGCGCGCAGGAGAGCGGCCGGCGTCAGCTCCTCGAC
CCCCGTGTCGGGCTAGTCCAGCGAGGCGACGGCGCGTGGGCCATGCCAGGCCGGCATGG
AGCGGTGGCGCGACCGCTGGCGCTGGTACGGGGCCTCGGGGGCATCGGCGGGCGTGGCC
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGCTGCGCCCGACTGTGGCAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTCAGGCTACCCGGGACTTGTACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC
TGCATCAACAATGCTGGCTTGGCCGGCCTGACACCCCTGCTCTCAGGCAGCACCGAGTGGTGGAA
GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGAAAGCCTACCAGTCCATGA
AGGAGCGGAATGTGGACGATGGCACATCATTAACATCAATAGCATGTCGGCCACCGAGTGTGTTA
CCCCCTGTCGTGACCCACTTCTATAGGCCACCAAGTATGCCGTACTGCCGTGACAGAGGGACT
GAGGCAAGAGCTCGGGAGGCCAGACCCACATCCGAGGCCACGTGCATCTCCAGGTGTGGTGG
AGACACAATTGCCCTCAAACCTCCACGACAAGGACCTGAGAAGGCAGCTGCCACCTATGAGCAA
ATGAAGTGTCTCAAACCGAGGATGTGGCGAGGCTGTATCTACGTCCCTCAGCACCCCCGGACA
CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACCTAGTGACTGTGGAGCTCC
TCCTTCCCTCCCCACCCCTCATGGCTTGCCTCCTGCCTCTGGATTTAGGTGTTGATTCTGGAT
CACGGGATACCACTCCTGTCCACACCCCGACCAGGGCTAGAAAATTGTTGAGATTTTATA
TCATCTGTCAAATTGCTTCAGTTGTAATGTGAAAATGGCTGGGAAAGGAGGTGGTGTCCC
TAATTGTTTACTTGTAACTTGTCTTGCCCTGGCACTTGGCTTGTCTGCTCTCAGTG
TCTTCCCTTGACATGGAAAGGAGTTGTGGCAAATCCCCATCTTGTGCACCTCAACGTCTG
TGGCTCAGGGCTGGGTGGCAGAGGGAGGCCACCTTATCTGTGTTGTTATCCAGGGCTCC
AGACTTCCCTCTGCCTGCCCACTGCACCCCTCCCCCTATCTATCTCCTCTCGGCTCCCC
AGCCCAGTCTGGCTTCTGTCCCTCCTGGGTACCCCTCAGTGTACTGACTATGGCAG
CAGAACACCAGGGCCTGGCCAGTGGATTCATGGTGCATTTAAGAAAAATCGCAACCAA
AAAAAAAAAA

FIGURE 128

MARPGMERWRDRLALVTGASGGIGAAVARALVQQGLKVVGCARTVGNIEELAAECKSAGYPGTLI
PYRCDLNEEDILSMFSAIRSQHSGVDICINNAGLARPDTLLSGSTSGWKDMFNVNLALSICTR
EAYQSMKERNVDDGHIININMSGHRVLPLSVTHFYSATKYAVTALTEGLRQEILREAQTHIRATC
ISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACTTCTACATGGGCCTCCTGCTGGTGCTTCCCTCAGCCTCCGCCGGTGGCCTACACCAT
CATGTCCCTCCCACCCCTTGACTGCGGGCCGTTCAGGTGAGAGTCTCAGTTGCCGGGAGC
ACCTCCCTCCCAGGGCAGTCTGCTCAGAGGGCCTCGGCCAGAATTCCAGTTGGTTCATGC
CAGCCTGTAAAAGGCCATGGAACTTGGGTGAATCACCGTGCCATTAAAGAGGGTTTCTGCCA
GGATGGAAATGTTAGGTCGTCTGTCTGCGTGTCATTCAGTAGCCACCAGCCACCTGTGG
CCGTTGAGTGCTTGAATGAGGAACTGAGAAAATTAATTCTCATGTATTTCTCATTTATTAA
TTATTTTAACTGATAGTTGACATATGGGGTACATGTGATTTGGAATGTTACAA
TATATAATGATCAAATCAGGGTAACTGGGATCCATCACATCAAACATTATTTATTCTT
TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTACTGCAAC
CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCTCCACCTCCAAGTAGCTGGACTACAGGCAT
GCACCACAAATGCCCAACTATTTGTTTAGAGAGACGGGTTTGCCATGTTGCCAGG
CTGGCCTTGAACTCCTGGCCTCAAACATCCACTGCCTCGGCCTCCAAGTGTTATGATTA
GGCGTGAGCCACCGTGCCTGGCTAAACATTTCTTGTGGGAACTTTGAATTAT
ACAATGAATTTTGTTACTGCTCCTGTGTGTGGTGGACTTGAATTT
ATCTAACTGTTTTGTACCAGTTAACCACCGTACTTCATCCCCACTCCCTCTATCC
AACCTCTGATCACCTCATTCTACTCTACCTCCATGAGATCCACTTTTTAGCTCCCACATGTG
AGTAAGAAAATGCAATTTGTTTCTGCCTGGCTTTAACTACATAATGACTCC
TTCCATCCATGTTGCGTGCAAATGACAGGATTCGTTCTTAATTCAATTAAAATACCACACATG
GCAAAAA

FIGURE 130

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSPRGSSLRGPRPRIPVLVSCQPV
KGHGTLGESPMFKRVFCQDGTVRSFCVCAVHFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTTGTATAAAGACCTAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
ATCTTCCTCATCGGGACTAAAATTGGGCTGTCCTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCCTG
TCCATCTGTGTCGCTGCGATGGGGTTTCATTTACTGTAATGATCGCTTCTGACATCCATTCCAACAG
GAATACCAGAGGATGCTACAACTCTACCTTCAGAACAAACAAATAATGCTGGATTCCCTTCAGAT
TTGAAAACCTGCTGAAAGTAGAAAGAATATACTATACCAACAGTTAGATGAATTCCCTACCAACCT
CCCAAAGTATGTAAGAGTACATTTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTCAA
AAATTCCCTATCTGGAAGAATTACATTAGATGACAACCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCA
TTCCGAGACAGCAACTATCTCGACTGCTTTCTGTCGGTAATCACCTAGCACAATTCCCTGGGGTT
GCCAGGACTATAGAAGAACTACGCTGGATGATAATCGCATATCCACTATTCATCACCACCTCTTCAG
GTCTCACTAGTCTAAACGCCTGGTCTAGATGAAACCTGTTGAACAATCATGGTTAGGTGACAAGTT
TTCTTCAACCTAGTTAATTGACAGAGCTGTCGGTGCAGAATTCCCTGACTGCTGCAACAGTAAACCT
TCCAGGCACAAACCTGAGGAAGCTTATCTCAAGATAACCACATCAATCGGGTGCCTTAAATGCTTTT
CTTATCTAAGGCAGCTCTATCGACTGGATATGTCCTAAATAAACCTAAGTAATTACCTCAGGGTATCTT
GATGATTGGACAATATAACACAACGTGATTCTCGAACAACTCCCTGGTATTGGGGTGCAGATGAAATG
GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGTGGGCTCATGTGCCAACCCCCAGAAA
AGGTTGGGGATGGCTATTAGGATCTCAATGCAAGACTGTTGATTGTAAGGACAGTGGGATTGTAAGC
ACCATTCAAGATAACCACGTCAATACCCAAACACAGTGTATCCGCCAACGGACAGTGGCCAGCTCCAGTGAC
CAAACAGCCAGATAATTAGAACCCCAAGCTCACTAAGGATCAACAAACCCACAGGGAGTCCCTCAAGAAAAAA
CAATTACAATTACTGTGAAGTGTCAACCTGTGACATACCATTCTATCTCTGGAAACTGCTCTACCTATG
ACTGCTTGAGACTCAGCTGGCTTAAACTGGGCATAGCCCCGGCATTTGGATCTATAACAGAAACAATTGT
AACAGGGGAAACGCAGTGAGTACTGGTCACAGCCCTGGAGCCTGATTCCCTATAAAGTATGCATGGTTC
CCATGGGAAACCCAGCAACCTCTACCTTGTGATGAAACTCCTGTTGATTGAGACTGAAACTGCACCCCTT
CGAATGTACAACCCCTACAACCCACCTCAATCGAGAGCAAGAGAAAAGAACCTTACAAAACCCAAATTACC
TTTGGCTGCCATCTGGTGGGGCTGTGGCCCTGGTTACCATGGCCCTTCTGCTTTAGTGTGTTGGTATG
TTCATAGGAATGGATCGCTTCTCAAGGAACGTGCAATATGCAAAGGGAGGAGAAGAAAGGATGACTAT
GCAGAAGCTGGCACTAAGAAGGACAACCTATCCTGGAAATCAGGAAACCTTCTTCAAGATGTTACCAAT
AAGCAATGAACCCATCTGAAGGGAGTTGTAATACACACCATACTTCTCCTAATGGAATGAATCTGT
ACAAAACAATCACAGTGAAGGAGCTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
TCACACTCATGATGCTGAAGGACTCACAGCAGACTGTGTTGGGTTTTAACCTAAGGGAGGTGATG
GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVRCDAAGFIYCNDRFLTSIPTGIPEDATTLYL
QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKEHLQENNIRTITYDSLSKIPYL
EELHLDNSVSAVSIEEGAFRDSNYLRLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL
QGLTSLKLRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQDNHIN
RVPNAFSYLRQLYRLDMSNNNLSNLPQGI FDDLDNITQLILRNNPWYCGCKMKWVWDWLQSLPV
KVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAQGQWPAPVTKQPD
IKNPKLTKDQQTTGSPSRKTITITVKSVTSDTIHISWKALPMTALRLSWLKLGHSPAFGSITET
IVTGERSEYLVTALEPDSPYKVCVPMETSNLYLFDETPVCIETETAPLRMYNPTTLNREQEKE
PYKNPNLPLAAIIGGAVALVTIALLALVCWYVHRNGSLFSRNCAYSKGRRKDDYAEAGTKKDNS
ILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNLKNNHSESSSNRSYRDSGIPDSDHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCCTCCCAGAGTCCTTGCCAGGCCACCCAGGCTTCTGGCA
GCCCTGCCGGGCCACTTGTCTTCATGTTCTGCCAGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG
TGCAGAGGCAGTCTGGCTTGGCCAGAGCTCAGGGTGTGAGCGTGTGACCAGCAGTGAGCAGAG
GCCGGCCATGGCCAGCCTGGGCTGCTGCTCTGCTTACTGACAGCACTGCCACCGCTGTGGT
CCTCCTCACTGCCTGGCTGGACACTGCTGAAAGTAAAGCCACCATGTCAGACACTGATCCTGTCT
GCGCTGGAGAGAGCCACCGTCTCCTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT
GGGGGTCGAGTGTGGAAGAGCAGCTAAAAAGTGTCCGGAGAAGTGGGCCAGGAGCCCTGC
TGCAGCCGCTGAGCCTGCGCTGGGATGCTGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
CTCCACTACCTCAAGCTGAGTGTGATCCCAAGTACCTAACAGAGAGTCCAGCTGACCCCTCAGCCCG
GTTTGAAAGCTCCCACATGCCCTGGATCCACACTGATGCCCTTGGTGTACCCACGTTGGC
CCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCTGGTGCAGCTGCTGGAAACGGGACG
GACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCGGCTGCTC
AGGCTACTGCCTGTCACCAACTGCTCTTCTCTGGGAGAATGAGGGATGCACACAGG
GACCACCCAACAGAGCCAGGACTATATCAACCTCTTGTGCGCCAACATGATGGACTTGAACCGC
AGAGCTGAGGCCATCGGATACGCCAACCTACCCGGACATCTCATGGAAAACATCATGTTCTG
TGGAAATGGGCGGCTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA
AACAGCAGGAAGGATGCTTCGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
TATCAGCAGCATTTCGAGGAGGTGAAGAGGCAGAAAAAACATTTCCAGATTCTCGCTCTGT
TGCTCAGGCTGGAGTACAGTGGCGCAATCTGGCTCACTGCAACCTTGCCTCTGGTTCAAGC
AATTCTTGCCTCATCCTCCGAGTAGCTGGACTACAGGAGCGTGCACCACACGGCTAAT
TTTATATTTTTAGTAGAGACAGGGTTCATCATGTTGCTCATGCTGGTCTCGAACTCTGAT
CTCAAGAGATCCGCCACCTCAGGCTCCAAAGTGTGGATTTAGGTGTGAGCCACCGTGTCTG
GCTGAAAAGCACTTCAAAGAGACTGTGTTGAATAAAGGGCCAAGGTTCTGCCACCCAGCACTC
ATGGGGCTCTCCCCTAGATGGCTGCTCTCCCACACACAGCCACAGCAGTGGCAGCCCTGG
GTGGCTCCTATACATCCTGGCAGAATACCCCCCAGCAAACAGAGGCCACACCCATCCACACCG
CCACCAAGCAGCGCTGAGACGGACGGTCCATGCCAGCTGCCAGGAGGAACAGACCC
TTAGTCCTCATCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG
ATAAGCAAAGCCACCCGACACCCAAATCTTGAAGCCCTGAGTAGGCAGGGCAGGGTAGGTGGG
GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAGTCAACTGCAACTGAAAAA
AAAAA

FIGURE 134

MSARGRWEGGRRACRGSILGLARAQGAERVTSSEQRPMASLGLLLLLTLAPPLWSSSLPGLD
TAESKATIADLILSALERATVFQRLPEINLDGMVGVRLLEEQLKSVREKWAQEPLLQPLSLRV
GMLGEKLEAAIQRSILHYLKLSDPKYLREFQTLTQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE
RSDVCLVQQLLGTGTDSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPLQQSQD
YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAIIWSWQKQQEGCFG
EPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFQFSCLILP
SSWDYRSVPYPLANFYIFLVETGFHHVAHAGLELLISRDPPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGGCCGCTCTGTGGGGCTTCTTCCGTCCTGCTGCTG
CTGCTGCTATCGGGGATGTCCAGAGCTGGAGGTGCCGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGGTCGGCATAGGAGATCGCTCAAGAGATTGAGGGCGTGCAGTTCCAGGGGTGAAGC
CTCAGGACTGGATCTCGCGGCCGAGTGCTGGTAGACGGAGAAGAGCACGTCGGTTCTTAAG
ACAGATGGGAGTTTGTGGTTCATGATATACTTCTGGATCTTATGTAGTGGAAAGTTGTATCTCC
AGCTTACAGATTGATCCCGTTGAGTGGATATCACTTCGAAAGGAAAATGAGAGCAAGATATG
TGAATTACATCAAACATCAGAGGTTGTCAAGACTGCCCTATCCTCTCCAAATGAAATCTTCAGGT
CCACCTCTTACTTTATTAAAAGGAATCGTGGGCTGGACAGACTTCTAATGAACCCAATGGT
TATGATGATGGTTCTCCCTTATTGATATTGTGCTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGAAATGGAGCAGTCAATGAATATGCTGAATTCCAACCATGAGTTGCCTGAT
GTTTCTGAGTTCATGACAAGACTCTCTCTCAAATCATCTGGAAATCTAGCAGCGCAGCAG
TAAAACAGGCAAAAGTGGGCTGGCAAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTGAC
AAACACGGCAACACTGGTGGCATCCAAGTCTGGAAAACCGTGTGAAGCAACTACTATAAAACTT
GAGTCATCCCGACGTTGATCTCTTACAACGTGTATGTT
AACTTTTAGCACATGTTGTACTTGGTACACGAGAAAACCCAGCTTCATCTTGTCTGTAT
GAGGTCAATATTGATGTCACTGAATTAATTACAGTGTCTATAGAAAATGCCATTAATAAAATTAT
ATGAACTACTATACTATTATGTATATTAATTAATTAACATCTTAATCCAGAAATCAAAAAAAA
AAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLSDVQSSEVPGAAEGSGSGVGIGDRFKIEGRAVPGVKPQDWISAA
RVLVDGEEHVGFLLKTDGSFVVHDIPSGSYVVEVVSPAYRFDPVRVDITSKGKMRARYVNYIKTSE
VVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPMMVLPLLIFVLLPKVVNTSDPDMRREME
QSMNMLNSNHELPDVSEFMTRLFSSKSSGKSSGSSKTGKSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTCTGTGAGATTGATTCTCCCCAGTCCCCGTGGGCTGAGGGGA
CCAGAAGGGTGGCTACGTTGGCTTCTGGAAAGGGGAGGCTATATGCGTCAATTCCCCAAACAA
GTTTGACATTCCCTGAAATGTCATTCTATCTATTCACTGCAAGTGCTGCTGTTCCAGGC
CTTACCTGCTGGCACTAACGGCGAGCCAGGATGGGACAGAATAAAGGAGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACCTCAGACCTGAAATCTTCTTCACGGGAGGCTTGGCAGT
TTTCTTACTCCTGTGGCTCCAGATTCAGGCCTAAGATGAAAGCCTCTAGTCTGCCTTCAGC
CTTCTCTGCTGCGTTTATCCTATGGACTCCTCCACTGGACTGAAGACACTCAATTGGG
AAGCTGTGTGATGCCACAAACCTCAGGAAATACGAAATGGATTTCTGAGATAACGGGCAGTG
TGCAAGCCAAGATGAAACATTGACATCAGAACTTAAAGGAGGACTGAGTCTTGCAAGACACA
AAGCCTGCGAATCGATGCTGCCCTGCCATTGCTAAGACTCTATCTGGACAGGGTATTAA
AAACTACCAGACCCCTGACCATTATACTCTCGGAAGATCAGCAGCCTGCCATTCCCTTCTTA
CCATCAAGAAGGACCTCCGGCTCTCATGCCACATGACATGCCATTGTGGGAGGAAGCAATG
AAGAAATACAGCCAGATTGAGTCACTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
TTTGGGAACTAGACATTCTCTGCAATGGATGGAGGAGACAGAAATGGAGGAAAGTGTGATGCTG
CTGCTAAGAATATTGAGGTCAAGAGCTCCAGTCTCAATACTGCAGAGGAGCATGCCCAA
ACCAACCCTCTTACTGTACTAGTCTTGTGCTGGTCACAGTGTATCTTATTTATGCATTACTTG
CTTCCCTGCAATTGCTTATGCATCCCCATCTAATTGAGACCATACTTGATAAGATTT
TGTAATATCTTCTGCTATTGGATATATTATTAGTTAATATATTATTTATTTGCTATTAA
ATGTATTTATTTTACTGGACATGAAACTTAAAAAAATTACAGATTATTTATAACCTG
ACTAGAGCAGGTGATGTATTTTATACAGTAAAAAAACCTTGAAATTCTAGAAGAGTGG
CTAGGGGGTTATTCAATTGTACTAAGGACATATTACTCATGCTGATGCTCTGTGAGAT
ATTGAAATTGAACCAATGACTACTTAGGATGGGTGGAATAAGTTGATGGAATTGCAC
ATCTACCTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATATTGTGTATCTCCAG
CCAGGAATCCTACACGGCCAGCATGTATTCTACAAATAAGTTCTTGCATACCAAAAAAA
AAAAAA

FIGURE 138

MRQFPKTSFDISPMSFSIYSLQVPAVPGLTCWALTAEPGWGQNKGATTCATNSHSDSELRPEIF
SSREAWQFFLLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTK PANRCCLLRHLLRLYLDdrvFKNYQTPDHYTLRKIS
SLANSFLTICKDLRLSHAHMTCHC GEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CTGGAGCCGGAAGCGCGCTGCAGCAGGGCAGGGCTCCAGGTGGGTCGGTCCGCATCCAGCC
TAGCGTGTCCACGATCGGGCTGGCTCCGGACTTCGCTACCTGTGGTAGCGATCGAGGTGC
TAGGGATCGCGGTCTTCCTCGGGGATTCTTCCCGGCTCCCGTCTGGAGCCAGTTCAACTGGACCACGCTGCC
CACGGAGCGGAGCCCCCAGCGCCGAACCCCTCGGCTGGAGCCAGTTCAACTGGACCACGCTGCC
ACCACCTCTCTCAGTAAAGTGTATTGTCTGATAGATGCCTGAGAGATGATTGTGTTG
GGTCAAAGGGTGTGAAATTATGCCCTACACAACCTACCTACCTGTGGAAAAGGAGCATCTCACAGT
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAATCAAGGCATTGATGACGGGGAG
CCTTCCTGGCTTGTGACGTCATCAGGAACCTCAATTCTCCTGCACTGCTGGAAAGACAGTGTGA
TAAGACAAGCAAAGCAGCTGGAAAAGAATAGTCTTTATGGAGATGAAACCTGGGTTAAATTA
TTCCCAAAGCATTGTGGAATATGATGGAACAACCTCATTTCGTCAGATTACACAGAGGT
GGATAATAATGTACGAGGCATTGGATAAAAGTATTAAAAAGAGGAGATTGGACATATTAATCC
TCCACTACCTGGGGCTGGACCACATTGGCCACATTTCAGGGCCCAACAGCCCCCTGATTGGCAG
AAGCTGAGCGAGATGGACAGCGCTGCTGATGAAAGATCCACACCTCACTGCACTGCGAAGGAGAGAGA
GACGCCCTTACCCAAATTGCTGGTTCTTGTGGTACCATGCTGAACAGGAAGTCACG
GGGCCTCCTCCACCGAGGGAGGTGAATACACCTCTGATTAACTCAGTTCTGCGTTGAAAGGAAA
CCCGGTGATATCCGACATCCAAAGCACGTCCTAGACGGATGTGGCTGCGACACTGGCAGTAGC
ACTTGGCTTACCGATTCCAAAAGACAGTGTAGGGAGCCTCTGCTTCCAGTTGTGGAAGGAAGAC
CAATGAGAGAGCAGTTGAGATTTCACATTGATACTAGTGCAGCTTAGTAAACTGTTGCAAGAG
AATGTGCCGTATGAAAAGATCCTGGTTGAGCAGTTAAATGTCAGAAAGATTGCAATGG
GAACCTGGATCAGACTGTACTTGGAGAAAAGCATTCAAGTCCTATTCAACCTGGCTCCAAGG
TTCTCAGGCAGTACCTGGATGCTCTGAAAGACGCTGAGCTTGTCCCTGAGTGACACAAGTGGCCAG
TTCTCACCCCTGCTCTGCTCAGCGTCCCACAGGCAGTGACAGAAAAGGCTGAGCTGGAAGTCCCC
CTGTCATCTCCTGGTTTCTGCTCTTATTTGGTGAATCTGTTCTTCCGGCCGTTACGT
CATTGTGTGCACCTCAGCTGAAAGTTCGCTACTTCTGTGGCTCTCGTGGCTGGCGAGGCT
GCCTTCGTTACCAAGACTCTGGTTGAACACCTGGTGTGCAAGTGTGGCAGTGCCCTGGAC
AGGGGCTCAGGGAAAGGACGTGGAGCAGCCTTATCCAGGGCTCTGGGTGTCCCACACAGGTG
TTCACATCTGTGCTGTCAGGTCAAGTCAGTTCTGGAAAGCTAGGTTCTGCGACTGTTAC
CAAGGTATTGTAAGAGACTGGCGGTCAAGAGGAACAAGCCCCCAGCTGAGGGAAAGAGACAATCGGCCTGGA
TCGGACAGCCTCCACAGAGGGTCAAAAGGAGACTTGGTCGACCAACTCATCCTGCCACCCAGAATGCATCCT
GCCTCATCAGGTCCAGATTCTTCCAAGCGGACGTTTCTGTTGAATTCTAGTCCTGGCC
TCGGACACCTTCATTGTTAGCTGGGAGTGGTGGTAGGCAGTGAAAGAGGGCGGATGGTCAC
ACTCAGATCCACAGAGGCCAGGATCAAGGGACCCACTGCACTGGCAGCAGGACTGTTGGCC
ACCCCAACCCCTGCACAGCCCTCATCCCTCTGGCTTGAGCCGTAGAGGCCCTGTGCTGAGTGT
CTGACCGAGACACTCACAGCTTGTCACTCAGGGCACAGGCTCCTCGGAGCCAGGATGATCTGTG
CCACGCTTGCACCTCGGGCCCATCTGGCTCATGCTCTCTCCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTAGTTACCAAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTLPPPLF
SKVVIVLIDALRDDFVFGSKGVKFMPYTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDSVIRQAKAAGKRIVFYGDETWKLFPKHFVEYDGTTSFVSDYTEVDNNV
TRHLDKVLKRGDWDLILHHLGLDHIGHISGPNSPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGGTATCGTACGCACCTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGTTGGAGTACAAC
TCCCTATAGAAAACA
CTGCCAGCACCTTAAGACCACTCACACCTTCAGAGTGAAGAACTTAAAC
CCGAAGAAATTCA
GCATT
CATGACCAGGATC
ACAAAGTACTGGTCTGGACTCTGGGAATCTCAT
AGCAGTTCCAGATA
AAAAACTACATACGCCAGAGATCTTCTTGCA
TTAGCCTCATCCTTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGTCTCTAAAGGGGAGTTTGTC
TAC
TGTGACAAGGATA
AAAGGACAAAGTC
ATCC
ATCC
TCAGCTGAAGAAGGAGAA
ACTGATGAAGCT
GGCTGCC
AAAGGAATCAGCACGCCGCC
CTTCATCTT
TATAGGGCTCAGGTGGCTCCTGGA
ACATGCTGGAGTC
GGCGCTCACCC
GGATGGTT
CATCTGC
CACCT
CTGCA
ATTGTA
ATGAGCCT
GTTGGGT
GACAGATA
AAATTG
AGAACAGG
AACACATTG
AATT
CAACCAG
TTGCAA
AGCTG
AAATG
AGGCC
AGT
GAGGT
CAGCGAT
TAGGAAACTGCC
CATTGA
AACGCC
TTCC
CTCG
CTA
ATTG
AACTA
ATTG
ATA
AAA
AC
CCA
AA
AC
CTG
CT
CA
CT
ACT

FIGURE 142

MLLLLEYNFPIENNQHLKTTHTFRVKNLPKKFSIHQDHKVLVLDGNLIAVPDKNYIRPEI
FFALASSLSSASAEGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAQKESARRPFI
FYRAQVGSWNMLESAAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36

N-myristoylation site.
amino acids 50-55, 87-92

Interleukin-1
amino acids 37-182

FIGURE 143

CTAGAGAGTATAGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTTAA
TCCAGGATCCTGTCCTCCTGTCCTGTAGGAGTGCCTGTCAGTGTTGGGTGAGACAAGTTG
TCCCACAGGGCTGTCTGAGCAGATAAGATAAAGGGCTGGGTCTGTGCTCAATTAACTCCTGTGG
CACGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCAATGCTGGCCTGCCGTGG
AAGGGAGGTCTGTCTGGCGCTGCTGCTGCTCTCTTAGGCTCCAGATCCTGCTGATCTATGC
CTGGCATTCCACGAGCAAAGGGACTGTGATGAACACAATGTCATGGCTCGTTACCTCCCTGCCA
CAGTGGAGTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCTACAGACTGGGG
CACATCTGAATTCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAACTAGGTGTGGAAATTGAAAGACGACATTGACAACGACTGCCATTCCAAGAAAGCACAG
AGCTGAACAATACTTCACCTGCTTCTTCACCATCAGCACCAAGGCCCTGGATGACTCAGTTCAGC
CTCCTGAACAAGACCTGCTGGAGGGATTCCACTTGAGTGAAACCCACTCACAGGCTTGTCCATGT
GCTGCTCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTCAGTGGCTGAGCAGCT
TTGGACTTGTGTTGATCCTATTGATGTGTTGAGATCTCAGATCAGTGTAGTTAGAAAATCC
ACACATCTTGAGCCTAATCATGTAGTGTAGATCATTAAACATCAGCATTAAAGAAAAAAAAAAAA
AAA

FIGURE 144

MLGLPWKGGLSWALLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP
WMTQFSLLNKTCLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCTGACGCG**GATGAGGAAG**
CACCTGAGCTGGTGGCTGGCACTGTCTGCATGCTGCTCTCAGCCACCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAATCAAGTGGAACCGGAAGGCCCTGCCAGCACTGCCAGATCA
CTGAGGCCAGGTGGCTGAGAACCGCCGGGAGCCTCATCAAGCAAGGCCAAGCTCGACATT
GACTTCGGAGGCCAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTCCCCGATGGCATCCA
CTACAAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTGTCACCGGCTGCATCAATGCCA
CCCAGGGCGCGAACCCAGGGGGAGTCCAGAACGCCAGACAACAGCTCCACCCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTGGTGGAGAGGGCGCAGGACT
TCGGGTACCATGCACCAGCCAGTGCCTCTGCCTTCTGGCTTGATCTGGCTCATGGT**GAAAT**
AAGCTTGCACGGAGGCTGGCAGTACAGAGGCCAGCGAGCAAATCCTGGCAAGTGACCCAGCT
CTTCTCCCCAAACCCACCGCTGGTCTGAAGGTGCCAGGAGCGCGATGCACTCGCACTGCAA
TGCGCTCCCACGTATGCGCCCTGGTATGTCCTGCCTGCTGATAAGATGGGGACTGTGGCTTCT
CCGTCACTCCATTCTCAGCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCAGTGCGCCACCTGCTCACAGTACTTCCAACAACTCTTAGAGGTAG
GTGTATTCCCGTTTACAGATAAGGAAACTGAGGCCAGAGAGCTGAAGTACTGCACCCAGCATC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCCCTGGCTTGCTAACCCCAAGGTTTCTGCTCT
GTCCAATTCCAGAGCTGTCTGGTATGTCACAGGGACCCACATCCAAACATGTAT
CTCTAAATGAAATTGTGAAAGCTCCATGTTAGAAATAATGAAAACACCTGA

FIGURE 146

MRKHLSWWWLATVCMLLFSHLSAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQGRK
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNKLHQO
VLWRLVQELCSLKHCFWLERGAGLRTMHQPVLLCLIALIWLVMVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCAAAGGGCTGGGATTATAGCGTGACCACATGCTGGTCCAGAGTCTCATTT
CCTGATGATTTATAGACTCAAAGAAAACTCATGTTCAGAAGCTCTTCTCTGGCCTCCCTCT
CTGTCTTCTTCCCTCTTCTTCTTAAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATCTTCATTTGCTTGTCACTGGGGTAGGTCACTGAGTCTAGTTTTATTTTGAATT
CAACTTTCAGATTCAAGGGGTACATGTGAAGGTTGTTATGAGTATATTGCATGATGCTGAGG
TTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLFLLILISSIYSESCKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTCACAGGAACCTCAGCACCCACAGGGCGGACAGCGCTCCCTCTACCTGGAGACTTGAC
TCCCGCGCGCCCCAACCTGCTTATCCCTGACCGTCGAGTGTAGAGATCCTGCAGCCGCCAGTCC
CGGCCCCCTCTCCGCCACACCCACCCCTCTGGCTCTTCTGTTTACTCCTCTTTCTTCATTCTATA
ACAAAAGCTACAGCTCCAGGAGCCAGCGCCGGCTGTGACCCAAGCCAGCGTGGAAAGAATGGGTT
CCTCGGGACCGGCACTGGATTCTGGTGTAGTGTCCCATTCAAGCTTCCCCAACCTGGAGGAA
GCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCAGAAAGACCTTGAATGAACAGATTGCTGAA
GCAGAAGAAGACAAGATTAACATATCCTCCAGAAAACAAGCCAGGTAGAGCAACTATTCTT
TGTTGATAACTTGAACCTGCTAAAGCAATAACAGAAAAGGAAAAATTGAGAAAGAAAGACAATCTA
TAAGAAGCTCCCCACTTGATAATAAGTGAATGTGGAAAGATGTTGATTCAACCAAGAATCGAAAATG
ATCGATGATTATGACTCTACTAAAGAGTGGATGGATCATAAATTCAAGATGATCCAGATGGTCTCA
TCAACTAGACGGGACTCTTAACCGCTGAAGACATTGTCATAAAATCGTGCAGGATTATGAAG
AAAATGACAGAGCCGTGTTGACAAGATTGTTCTAAACTACTTAATCTGGCCTTATCACAGAAAGC
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTACAAAATTAACTCTCAAAGGAAGCCAACAA
TTATGAGGAGGATCCAATAAGCCCACAAGCTGGACTGAGAATCAGGCTGGAAAAATACCAGAGAAAG
TGACTCCAATGGCAGCAATTCAAGATGGCTTGCTAAGGGAGAAAAGATGAAACAGTATCTAACACA
TTAACCTTGACAAATGGCTGGAAAGGAGAACTAAACCTACAGTGAAGACAACCTTGAGGAACCTCA
ATATTTCCAAATTCTATGCGCTACTGAAAAGTATTGATTGAGCTAGAAAAAGCAAAAGAGAAAGAAA
CACTGATTACTATCATGAAAACACTGATTGACTTGTGAAGATGATGGTAAATATGGAACAATATCT
CCAGAAGAAGGTGTTCTACCTGAAAACCTGGATGAAATGATTGCTCTCAGACACAAAAACAAGCT
AGAAAAAAATGCTACTGACAATATAAGCAAGCTTCCAGCACCACAGAGAAGAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCCACAAAA
GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAACAGAACGCTATTGGAAGC
CATCAGAAAAATATTGAATGGTGAAGAAACATGACAAAAGGAAATAAGAAGATTATGACCTT
CAAAGATGAGAGACTTCATCAATAAACAAAGCTGATGCTTATGTGGAGAAAGGCATCCTGACAAGGAA
GAAGCCGAGGCCATCAAGCGATTAGCAGCCTGAAAAATGGCAAAGATCCAGGAGTCTTCAA
CTGTTTCAGAAAACATAATATAGCTTAAACACTCTAATTCTGTGATTAAAATTTTGACCCAAGG
GTTATTAGAAAGTGTGAATTACAGTAGTTAACCTTACAAGTGGTTAAACATAGCTTCTCCC
GTAAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAAAAAAA

FIGURE 150

MGFLGTGTWILVLVLPIQAFPKPGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTPPENKPG
QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLINVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDFKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEEDPNKPTSWTENQAGKIKEVKTPMAAIQDGLAKGENDETVSNTLTNTNGLE
RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV
SYLENLDEMIALQTKNLEKNATDNISKLFPAPEKSHEETDSTKEEAAKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKGKGNKEDYDLSKMRDFINKQADAYVEKGILD
EEAEAIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCGTGGAGCTCAA
GATGGTCTGAGTGGGGCGCTGTGCTCCGATGAAGGACTCGCATTGAAGGTGCTTATCTGC
ATAATAACCAGCTCTAGCTGGAGGGCTGCATGCAGGGAAAGTCATTAAGGTGAAGAGATCAGC
GTGGTCCCCAATCGTGGCTGGATGCCAGGCTGTCCCCCGTACCTGGGTGTCAGGGTGGAG
CCAGTGCCTGTCATGTGGGTGGGCAGGGCCACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTGGTGCAGGAACTCAAGAGCTTCACCTCTACGGCGGGACATGGGCTCACC
TCCAGCTCGAGTCGGCTGCCTACCCGGCTGGTCTGTGACGGTGCAGGGTGGAG
TGTCAACTCACCCAGCTCCCAGAGAATGGTGGCTGGAAATGCCCATCACAGACTCTACTTCC
AGCAGTGTGACTAGGGCAACGTGCCCCCAGAACCTCCCTGGGAGAGCCAGCTGGGTGAGGGT
GAGTGGAGGAGACCCATGGCGGACAATCACTCTCTGCTCTCAGGACCCCCACGTCTGACTTAG
TGGGCACCTGACCACTTGTCTCTGGTCCAGTTGGATAAATTCTGAGATTGGAGCTCAGT
CCACGGTCCCTCCCCACTGGATGGTGCTACTGCTGTGAACTTGTAAAAACCATGTGGGTAAA
CTGGGAATAACATGAAAAGATTCTGTGGGGTGGGGGGGGAGTGGTGGGAATATTCTGCT
TAATGGTAACTGACAAGTGTACCCCTGAGCCCGAGGCCAACCCATCCCCAGTTGAGCCTATA
GGGTCAGTAGCTCCACATGAAGTCTGCACTCACCCTGTGCAAGGAGAGGGAGGTGGTATA
GAGTCAGGGATCTATGGCCCTTGGCCAGCCCCACCCCCCTTCCCTTAATCCTGCCACTGTCTATA
TGCTACCTTCCATCTTCCCTCATCATCTTGTGTTGGGATGAGGAGGTGGTATGTGAGAA
GAAATGGCTGAGCTCAGAAGATAAAAGATAAGTAGGTATGCTGATCCTCTTTAAAAACCAA
GATACAATCAAATCCCAGATGCTGGTCTATTCCATGAAAAAGTGTCTGACATATTGAGA
AGACCTACTTACAAGTGGCATATATTGCAATTATTAAATTAAAAGATAACCTATTATATT
TCTTTATAGAAAAAGTCTGGAAGAGTTACTCAATTGTAGCAATGTCAAGGGTGGCAGTAT
AGGTGATTTTCTTTAATTCTGTTAATTATCTGTATTCTCTAATTCTACATGAAGATGA
ATTCCCTGTATAAAAATAAGAAAAGAAATTATCTGAGGTAAAGCAGAGCAGACATCATCTG
TTGTCCTCAGCCTCCACTTCCCAGAGTAAATTCAAATTGAATCGAGCTCTGCTCTGGTTGG
TTGTAGTAGTGTAGTCAAGGAAACAGATCTCAGCAAAGCCACTGAGGAGGAGGCTGTGCTGAGTTGT
GTGGCTGGAAATCTCTGGTAAGGAACCTAAAGAACAAAATCATCTGTAATTCTTCTAGAAG
GATCACAGCCCCCTGGGATTCCAAGGCATTGGATCCAGTCTAAGAAGGCTGCTGACTGGTTGA
ATTGTGTCCTCTCAAATTCACTACCTCTGGAAATCTCAGTCTGTGAGTTATTGGAGATAAG
GTCCTGCAGATGTAGTTAGTAAAGACAAGGTATGCTGGATGAAGGTAGACCTAAATTCAATAT
GACTGGTTCTGTATGAAAAGGGAGGAGGACAGAGACAGAGGAGACGCGGGGAAGACTATGTA
AAGATGAAGGCAGAGATCGGAGTTTGCAGCCACAAGCTAAGAAAACACCAAGGATTGTGCAACC
ATCAGAAGCTTGGAGAGGCAAGAAGAATTCTCCCTAGAGGCTTAGAGGGATAACGGCTCTG
CTGAAACCTTAATCTCAGACTTCCAGCCTCTGAACGAGAAAGAATAAATTCTGGCTGTTTAA
GCCACCAAGGATAATTGGTACAGCAGCTCTAGGAAACATAACAGCTGCTAAATGATCCTGT
CTCCTCGTGTTCACATTCTGTGTGTGCCCCCTCCACATGTACCAAAGTTGTCTTGTGACCAA
TAGAATATGGCAGAAGTGTGGCATGCCACTTCAAGAGATTAGTTATAAAAGACACTGCAGCTC
TACTTGAGCCCTCTCTCTGCCACCCACCGCCCCAATCTATCTGGCTCACTCGCTCTGGGG
AAGCTAGCTGCCATGCTATGAGCAGGCCATAAAAGAGACTTACGTGGTAAAAATGAAGTCTCCT
GCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATCGATTTGTTGTTTAA
AAGTTGCTCAGTTGGCTAACTGTTATGCAGCAATAGATAAAATATGCAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEELSVVPNRWLDASLSPVILGVQGGS
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
VRLTQLPENGGWNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCGC
CCTGCAGAAATCTGTGAGCTTTCTTATGGGGACCTGGCCACCAGCTGCCTCCTCTCTGG
CCCTCTGGTACAGGGAGGAGCAGCTGCGCCATCAGCTCCACTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATCACCAACCGCACCTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
CAACACAGACGTTCGTCTCATGGGAGAAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACCTCACCCCTGAAGAAAGTGCTGTTCCCTCAATCTGATAGGTT
CAGCCTTATATGCAGGAGGTGGTGCCTGGCCAGGCTCAGCAACAGGCTAACGACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTGCTGTTATGTCCTGAGAAAT
GCCTGCATTTGACCAAGAGCAAAGCTGAAAAATGAATAACTAACCCCCCTTCCCTGCTAGAAATAA
CAATTAGATGCCCAAAGCGATTTTTAACCAAAAGGAAGATGGGAAGCCAAACTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAACCAATGCCACTTTGTTATA
AGACCAGAAGGTAGACTTCTAAGCATAGATAATTATTGATAACATTCAATTGTAACTGGTGTTC
TATACACAGAAAACAATTATTAAATAATTGTCTTTCCATAAAAGATTACTTTCCAT
TCCTTACGGGAAAAACCCCTAAATAGCTCATGTTCCATAATCAGTACTTTATTTATAAA
TGTATTATTATTATTATAAGACTGCATTTATTATCATTATTAAATATGGATTATTAT
AGAAACATCATTGATATTGCTACTTGAGTGTAAGGCTAATATTGATATTATGACAATAATTAT
AGAGCTATAACATGTTATTGACCTCAATAAACACTGGATATCCC

FIGURE 154

MAALQKSVSSFLMGTLATSCLLLALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNNTDVRLLIGEKLFHGVSMERCYLMQVLNFTLEEVLFQSDRFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT
CAGTCAGTGCCCGACTTGTGACTGAGTGTGCAGTGCCAGCATGTACCAGGTAGTCAGAGGGC
TGCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCAAGC
TGCCAGGTTGGGGCTGGGGCCAAGTGGAGTGAGAAACTGGGATCCCAGGGGAGGGTGCAGAT
GAGGGAGCAGCCACAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTCTACAGGTGGTTGCAT
TCTTGGCAATGGTCATGGGAACCCACACCTACAGCCACTGGCCAGCTGCTGCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCCTCCCTAGAGCCTGCTAG
GCCCAACCGCCACCCAGAGTCCTGTAGGGCCAGTGAAGATGGACCCCTAACACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCCAGGACCTGTACCAAGCCCGT
TGCCTGTGCCCGACTGCGTCAGCCTACAGACAGGCTCCACATGGACCCCCGGGCAACTCGGA
GCTGCTCTACCACAAACCAACAGACTGTCTCTACAGGCGCCATGCCATGGCGAGAAGGGCACCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTAGCTTGTGTGTGCGGGCCCGT
GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTGGAAACCTGGAGCCAGGTGTACA
ACCACTGCCATGAAGGGCCAGGATGCCAGATGCTGGCCCTGTGAAGTGTCTGGAGCAG
CAGGATCCCAGGACAGGATGGGGCTTGGGAAACCTGCACTCTGCACATTGAAAAGAG
CAGCTGCTGCTAGGGCCGCCGGAAAGCTGGTGTCTGTCACTTCTCTCAGGAAAGGTTTCAA
GTTCTGCCATTCTGGAGGCCACCACTCCTGTCTCTCCCTCTTCCATCCCCGTACCCCTG
GCCCAGCACAGGCACATTCTAGATATTCCCCCTGCTGGAGAAGAAAGGCCCTGGTTTATT
TGTGACATGGATGATTCTGAGGAGGAGCTGTTATTGAATGTATAGAGATTATCAAATAAATAT
CTTTATTTAAAAATGAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGTHTYSHWPSCCPGQDTSEELLRWSTVPVPPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCG**ATG**TCGCTCGTGTGCTAACGCTGGCCGCGTGTGCAGGAGCGCCGTACCCGAGAGCC
GACCGTCAATGTGGCTCTGAAACTGGGCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACATAGTGTGCAACAGGGGACTATTCA
ATTTGATGAATGTAAGCTGGGTACTCCGGCAGATGCCAGCATCCGTTGTTGAAGGCCACCAA
GATTTGTGTGACGGGAAAGCAACTTCCAGTCTCACAGCTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCCTCTGGTGGTAAATGGACATTTCTACATCGGCTTCCCTGTA
GAGCTGAACACAGCTATTCTATTGGGCCATAATATTCTTAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTTCACCTCACCAAGGCTGCCTAGACCACATAATGAAATATAAAAAAA
AGTGTGTCAAGGCCGGAAAGCCTGTGGATCCGAACATCACTGCTGTAAGAAGAATGAGGAGACA
GTAGAACTGAACCAACCACCCCTGGGAAACAGATACTGGCTTATCCAACACAGCAC
TATCATGGGTTTCAGGTGTTGAGCCACACCAGAAAGAAACAAACGCGAGCTTCAGTGGTGA
TTCCAGTGAATGGGATAGTGAAGGTGCTACGGTCAGCTGACTCCATATTCTACTTGTGGC
AGCGACTGCATCCGACATAAGGAACAGTTGTGCTGCCCCACAAACAGGCGTCCCTTCCCT
GGATAACAAACAAAAGCAAGCCGGAGGCTGGCTGCCTCTCCTGCTGTCTGCTGGTGC
CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTTCTT
TCTACCACCAACTACTGCCCTTCAATTAGGTTCTGTGGTTTACCCATCTGAAATATGTTCCA
TCACACAATTGTTACTTCACTGAATTCTTCAAAACCATTCAGAAGTGGTCATCCTTGAAA
AGTGGCAGAAAAGAAAATAGCAGAGATGGGTCCAGTGCAGTGGCTGCCACTAAAAGAAGGCA
GCAGACAAAGTCGCTTCTTCAATTGACGTCAACAGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCAGTGAGAACTCTCAAGACCTCTCCCCCTGCCTTAACCTTCTGCA
GTGATCTAAGAAGCCAGATTCTGACAAATACGTGGTGGTCACTTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTCAGTGTCTGCCCAAGTACCACTCATGAAGGATGCCACTGCTT
CTGTGCAGAACATTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAGATCACAAGCCTGCCACG
ATGGCTGCTCCTTGT**AG**

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLIPGDLRDLRVEPVTSVATGDYSILMNVS梧
LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP
NANMNEDGPMNSVNFSPGCLDHIMKYKKKCVKAGSLWDPNITACKNEETVEVNFTTPIGNRYMALIQH
STIIGFSQVFEPHQKKQTRASVVIPTGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPQTGVPFPLDNK
SKPGGWLPLLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTLLPPIKVLVVYPSEICFHHTICYFTEFL
QNHCRCSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLSNDVNSVCDGTCGKSEGPSENSQDLFPLA
FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAELLHVVKQQVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristylation site.

amino acids 116-122

Amidation site.

amino acids 488-492

FIGURE 159

AGCCACCAAGCGCAACATGACAGTGAAGACCCATGCCAGCCATGGTCAAGTACTTGCTGCT
GTCGATATTGGGGCTTGCCTTCTGAGTGAGGCGCAGCTCGAAAATCCCCAAGTAGGACATA
CTTTTTCCAAAAGCTGAGAGTTGCCGCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCTTCCATGTCACGTAACATCGAGAGCCGCTCACCTCCCCCTG
GAATTACACTGTCACTGGGACCCCAACCGGTACCCCTCGGAAGTTGTACAGGCCAGTGCTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCATCCAGCAA
GAGACCCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTGTGTTCTTCCAGTTGGAGAAGGTGCT
GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAGTAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 16O

MTVKTLHGPAMVKYLLSILGLAFLSEAARKIPKGHTFFQKPESCPPPGGSMKLDIGIINEN
QRVMSMRNIESRSTSPWNYTVTWDPNRYPSEVVQAQCRLGCINAQGKEDISMNSVPIQQETLVV
RRKHQGCSVSFQLEKVLVTVGCTCVTPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGCCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGACTCAGGACTCCAGG
ACAGAGAGTGCACAAACTACCCAGCACAGCCCCCTCCGCCCCCTGGAGGCTGAAGAGGGATT
CAGCCCCCTGCCACCCACAGACACGGGCTGACTGGGGTGTCTGCCCCCTGGGGGGGGCAC
AGGGCCTCAGGCCTGGGTGCCACCTGGCACCTAGAAGATGCCTGTGCCCTGGTCTTGCTGC
TGGCACTGGGCCAGGCCAGTGGTCTTCTCTGGAGAGGCTGTGGGCCCTAGGACGCTACC
CACTGCTCTCCGGGCTCTCCTGCCCTCTGGACAGTGACATACTCTGCCTGCCTGGGACAT
CGTGCCTGCTCCGGGCCGTGCTGGCCTACGCACCTGCAGACAGAGCTGGTGTGAGGTGCC
AGAAGGAGACCGACTGTGACCTCTGTCTGCGTGTGGCTGTCCACTGGCGTGATGGGACTGG
GAAGAGCCTGAAGATGAGGAAAAGTTGGAGGAGCAGTGACTCAGGGTGGAGGAGCCTAGGAA
TGCCTCTCTCCAGGCCAAGTCGTGCTCCTCCAGGCCTACCCACTGCCCCGTGCGTCTGC
TGGAGGTGCAAGTGCCTGCTGCCCTGTGCACTTGTGCACTGTGGGCTCTGGTATATGAC
TGCTTCAGGGCTGCCCTAGGGAGTGAGGTAGAATCTGGCTTATACTCAGCCCAGGTACGAGAA
GGAACACTAACACACACAGCAGCTGCCCTGCCCTGGCTCAACGTGTCAAGCAGATGGTACA
ACGTGCATCTGGTCTGAATGTCTGAGGAGCAGACTTCGGCCTCTCCCTGTACTGGAATCAG
GTCCAGGGCCCCAAAACCCGGTGGCACAAAACCTGACTGGACCGCAGATATTACCTTGAA
CCACACAGACCTGGTCCCTGCTCTGTATTCAAGGTGTGGCCTCTGGAACCTGACTCCGTTAGGA
CGAACATCTGCCCTTCAGGGAGGACCCCCCGCGCACACCAGAACCTCTGGCAAGCCGCCGACTG
CGACTGCTGACCCCTGAGAGCTGGCTGGACGCACCGTGTGCGTCCCCGAGAACGGGACT
GTGCTGGCGGGCTCCGGTGGGACCCCTGCCAGCCACTGGTCCCACCGCTTCTGGAGAACG
TCACTGTGGACAAGGTTCTCGAGTCCCATTGCTGAAAGGCCACCTAACCTCTGTGTTCAAGGTG
AACAGCTGGAGAAGCTGCAGCTGCAGGAGTGTGCTGGGCTGACTCCCTGGGCCCTCTCAAAGA
CGATGTGCTACTGTGGAGACACGAGGGCCCCAGGACAACAGATCCCTCTGTGCCCTGGAACCCA
GTGGCTGTAACCTCACTACCCAGCAAGCCCTCACGAGGGCAGCTGCCCTGGAGAGTACTTACTA
CAAGACCTGCAGTCAGGCCAGTGTCTGCACTATGGGACGATGACTTGGAGGGCAGCTATGGGCTG
CCCCATGGACAAATACATCCACAAGCTGGGCTCTGGCTGTGGCTGCCCTGCCTACTCTTGCCG
CTGCGCTTCCCTCATCCTCTCAAAAGGATCACCGAAGGGTGGCTGAGGCTCTTGAA
CAGGACGCTCCGCTGGGGGGCGCCAGGGGCCGCGGCTCTGCTCTACTCAGCCGATG
CTCGGGTTTCAGGCGCTGGGGGCCCTGGCGTGGCCAGCTGCCGCTGCCGTGG
CCGTAGACCTGTGGAGCCGTGGTGAAGTGAAGCGCGCAGGGGCCGTTGGCTTCA
CGGCCAGACCTGAGGGCGCGTGGTGGCTTGTGCTCTCTCCCGGTGCCGTGG
GTGCAAGCGAGTGGCTACAGGATGGGTGTCCGGGCCGGCGCACGGCCG
GCCCTCGCTCAGCTCGTGTGCCCCACTTCTGCAAGGGCGGCCGGCAGCTACGTGGGG
GCCCTCGACAGGTGCTCCACCCGGACGCCGTACCCGCCCTTTCGCACCGTGCCGTCTT
CACACTGCCCTCCCAACTGCCAGACTTCTGGGGCCCTGCAAGCAGCCTGCCGCCG
GGCGGCCAGAGAGAGAGCGGGAGCAAGTGTCCGGGCCCTCAGGCCAGGCC
CATCCCCCGGGGACTCCCGCCGGGACGCCGGGTGGGACCGGGGCCG
CGGGACTAAATAAAGGCAGACGCTGTTTCTAAAAAA

FIGURE 162

MPVPWFLLSLALGRSPVVLSLERLVGPQDATHCSPGLSCRILWSDSILCLPGDIVPAPGPVLAPTHLQTELVLRCQKETDCDLCRLRVAVHLAVGHWEEPEDEEKFGGAADSGVEEPRNASIQAQVVLSFQAYPTARCVLLEVQVPAALVQFGQSVGSVYDCFEAALGSEVRIWSYTQPRYEKELNHTQQLPALPWLNVSADGDNVHLVLNVEEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQNLWQAARLRLLLTQSWSLLDAPCSLPAEAALCWRAPGGDPCQPLVPPPLSWENVTVDKYLEFPLLKGHPNLCVQVNSSEKLQLQECLWADSLGPLKDGVLLLETRGQPQDNRSILCALEPSGCTSILPSKASTRAARLGEYLLQDLQSQCQLQWDDDLGALWACPMDKYIHKRWALVWLACLLFAAALSLLLLKKDHAKGWLRLKQDVRSGAAARGRAALLLYSADDSGFERLVLGALASALCQLPLRVAVDLWSRRELSAQGPVAWFHAQRRQTLQEGGVVVLFSPGAVALCSEWLQDGVSQPGGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFRTPVFTLPSQLPDFLGALQQPRAPRSGRLQERAEQVSRALQPALDSYFHPPGTPAPGRGVGPGAGPGAGDGT

Signal sequence:
amino acids 1-20

Transmembrane domain.
amino acids 453-475

N-glycosylation sites.
amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251, 334-337, 357-360, 391-394

Glycosaminoglycan attachment site.
amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 552-555

N-myristoylation sites.
amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617, 692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTGACTGTGGGATCCCTGGCT
GCTCACGCCCTGAGGACCCCTGGATCTGCTCCAGCACGTGAAATTCCAGTCAGCAACTTGA
AAACATCCTGACGTGGGACAGCAGGGCAGAGGGCACCCAGACACGGTCTACAGCATTGAGATA
AGACGTACGGAGAGAGGGACTGGGATGGCAAGAAGGGTGTCACTGGGATCACCCGGAAAGTCTGC
AACCTGACGGTGGAGACGGCAACCTCACGGAGCTACTATGCCAGGGTCACCGCT
GTCAGTGCGGAGGCCGGTCAGCCACCAAGATGACTGACAGGTTCACTCTGCAGCACACTAC
CCTCAAGGCCACCTGATGTGACCTGTATCTCAAAGTGAGATGATTCACTGAGATGATTGTTCATCCTA
CCCCCCACGCCAACCGTGCAGCGATGCCACCGGCTAACCTGGAAGACATCTTCCATGACCTG
TTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACCTTGGGAGGGAAAGCAGAGAGA
ATATGAGTTCTTCCGGCTGACCCCTGACACAGAGTCTTGGGACCATCATGATTGCGTTCCA
CCTGGGCCAAGGAGAGTGCCCTTACATGTGCCAGTGAGACACTGCCAGACCGGACATGGACC
TACTCCTCTCCGGAGCCTCTGTTCTCCATGGGCTCCCTGTCGAGTACTCTGCTACCTGAG
CTACAGATATGTCAACCAAGGCCCTGCACCTCCAACTCCCTGAACGTCAGCGAGTCCTGACTT
TCCAGCCGCTGCCTTCACTCAGGAGCACGTCCTGATCCCTGTTGACCTCAGCGGGCCCCAGC
AGTCTGGCCCAGCCTGTCCAGTACTCCAGATCAGGGTGTCTGGACCCAGGGAGGCCGCAGGAGC
TCCACAGCGGCTAGCCTGTCCAGATCACCTACTTAGGGCAGCAGACATCTCATCCTCCAGC
CCTCAACGTGCCACCTCCCCAGATCCTCTCCCCACTGTCTATGCCCAAAAGCCTGCCCTGAG
GTCGGGCCCCATCCTATGCACCTCAGGTGACCCCCGAAGCTCAATTCCATTCTACGCCCA
GGCCATCTCTAAGGTCAGCCTCTCTATGCCCTCAAGCCACTCCGGACAGCTGGCCTCCCT
CCTATGGGTATGCAAGGTTCTGGCAAAGACTCCCCACTGGGACACTTTCTAGTCTAA
CACCTTAGGCCCTAAAGGTCAAGCTCAGAAAGGCCACAGCTGGAGCTGCATGTTAGGTGCCCT
TTCTCTGCAGGAGGTGACCTCTGGCTATGGAGGAATCCAAGAAGCAGAAATCATTGCACCCAGC
CCCTGGGGATTGACAGACAGAACATCTGACCCAAATGTGCTACACAGTGGGAGGAAGGGACA
CCACAGTACCTAAAGGCCAGCTCCCCCTCTCTCCAGTCCAGATCGAGGGCCACCCCATGTC
CCTCCCTTGCACACTCCTCCGGTCCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGGCC
TGCTGGAGTCCCTGTGTGTCCTGAAGGATGAAGCCAAGGCCAGGGCTGAGACCTCAGACCTG
GAGCAGCCCACAGAACACTGGATTCTCTTTCAAGGGCTGGCCCTGACTGTGCAAGGGAGTCCTG
AGGGGAATGGGAAAGGCTTGGCTTCTCCCTGCTCCCTACCCAGTGTACATCCTGGCTGTCA
ATCCCATGCCTGCCCATGCCACACACTCTGCGATCTGCCCTCAGACGGGTGCCCTTGAGAGAAC
AGAGGGAGTGGCATGCAGGGCCCTGCCATGGTGCGCTCTCACCGAACAAAGCAGCATGATA
AGGACTGCAGGGGGAGCTGGGAGCAGCTGTGTAGACAAGCGCTGCTCGTGAGCCCTG
CAAGGCAGAAATGACAGTGCAGGAGGAAATGCAGGGAAACTCCGAGGTCCAGAGCCCCACCTC
CTAACACCATGGATTCAAAGTGTCAAGGAATTGCTCTCCCTGCCCTTGTCAATTGTTAAAGGTGGGAAGAGA
ACATCTGAGAATGGGTTGAAAGGAAGGTGAGGGCTGTGGCCCTGGACGGGTACAATAACACAC
TGTACTGATGTCACAACCTTGTCAAGCTCTGCCCTGGGTCAGGCCATCTGGCTCAAATTCCAGC
CTCACCACTCACAAGCTGTGACTTCAAACAAATGAAATCAGTGCCAGAACCTCGGTTCTC
ATCTGTAATGTGGGATCATAACACCTACCTCATGGAGTTGTGGTGAAGATGAAATGAGTCATG
TCTTAAAGTGTAAATGTCATGGCTGGTACATGGCAGTGCCTAAACACGGTAGCTATTTAAAAA
AAAAAAA

FIGURE 164

MRTLLTILTGVSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTVYSIEYKTYGERDW
VAKKGCQRITRKSCNLTVETGNLTLYYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVTCIS
KVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHLGGKQREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCRVKTLPDRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP
PNSLNQVRVLTFOPLRFIQEHVLIPIVFDLSPGSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFPFYAPQAISKVQPSSY
APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPKGQLQKEPPAGSCMLGLSLQEVTSIAM
EESQEAKSLHQPLGICTDRTSDPNVLHSGEEGTPQYLGQLPLLSSVQIEGHPMSLPLQPPSGPC
SPSDQGPSPWGLLESLVCPKDEAKSPAPETSDLEQPTELDSLFRGLALTQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAGTCACCCGGGCCGCGGTGCCACAACATGG
CTGGCGGCCGGGCTGCTCTGGCTGTCGTGCTGGGGCGCTCTGGGGTCCCAGGCCAG
TCGGATCTCAGCCACGGACGGCGTTCTGGACCTCAAAGTGTGCGGGGACGAAGAGTGCAGCAT
GTTAATGTACCGTGGAAAGCTTGAAGACTCACGGCCCTGATTGTCGTTTGTGAATTAA
AAAAAGGTGACGATGTATATGTTACTACAAACTGGCAGGGGATCCCTGAACTTGGCTGGA
AGTGTGAACACAGTTGGATATTTCAAAGATTGATCAAGGTACTTCATAAATACACGGA
AGAAGAGCTACATATTCCAGCAGATGAGACAGACTTTGTCGCTTGAAGGAGGAAGAGATGATT
TTAATAGTTATAATGAGAGCTTTAGGATCTTGGAACTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAAGTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGGCGTGA
ACTTGACCTGTGCCTGAGCCCCAGGCATTCAAGAGCTGATTCAAGGGATGGAGAAGGTGCTTCT
CAGAGAGCACCGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCCACACCGCGT
CCTGCGGCTAACGCTCAGGGAGTGCAGTCTCGTGGACACTTTGAAGAAATTCTGCACGATAA
ATTGAAAGTGCCAGGAAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGTGGAGCGGGAGA
AGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCCTTATT
CATTACAGCAAAGGATTCGTTGGCATCAAAATCTAAGTTGTTTACAAAGATTGTTTAGTA
CTAAGCTGCCTGGCAGTTGCATTTGAGCAAACAAAAATATTATTTCCCTCTAAGTA
AAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLFWLFVLGALWWVPGQSDLHGRRFSDLKVCGLDEECMILMYRGKALEDFTGPDCRFVN
FKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEEELHI PADETDFVCFEGGRD
DFNSYNVEELLGSLELEDSPVEESKKAAEVSQHREKSPPEESRGRELDPVPEAFRADSEDEGGA
FSESTEGLQGQPSAQESHPTSGPAANAQGVQSSLDTFEEILHDKLKVGSESRTGNSSPASVER
EKTDAYKVLKTEMSQRGSGQCVIHYSKGFRWHQNLSLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTCAAGAGGCCGGGAAGAGAAGCAAAGCGC
AACGGTGTGGTCCAAGCCGGGCTTCTGCTCGCTCTAGGACATAACACGGGACCCCTAACTTC
AGTCCCCCAAACGCGCACCTCGAAGTCTTGAACCTCAGCCCCGACATCCACCGCGGCCACAGG
CGCGGCAGGCAGGTCCCGGCCGAAGGCAGTCGCGCAGGGGTCGGGCAGCTGGGCTCGGGC
GGCAGGGAGTAGGGCCCGCAGGGAGGAGGGCTGCATATTCAAGAGTCGCGGGCTCGCCCTG
GGCAGAGGCCGCCCTCGCTCACGCAACACCTGCTGCTGCCACCGCGCCGCGATGAGGCCGTGG
TCTCGCTGCTGCTGGCGCCCGCTGCTCTGCGGCCACGGAGCCTCTGCCGCCGCGTGGTCAGC
GCCAAAAGGTGTGTTGCTGACTTCAGCAGTACAAATGGCCTACTTCCATGAAC
GTCCAGCCGAGTGAAGCTTCAGGAGGCACGCCCTGGCTGTGAGAGTGAGGGAGGTCTCCTCA
GCCTTGAGAATGAAGCAGAACAGAACAGTTAATAGAGAGCATGTTGAAACACCTGACAAAACCGGG
ACAGGGATTCTGATGGTATTCTGGATAGGGCTTGGAGGAATGGAGATGGCAAACATCTGG
TGCCTGCCAGATCTTACCAAGTGGCTGATGGAAGCAATTCCAGTACCGAAACTGGTACACAG
ATGAACCTCCTCGCGGAAGTGAAGAAGTGTGTTGATGTATCACCAACCAACTGCCAACCTGGC
CTTGGGGGTCCCTACCTTACCAAGTGGATGACAGGTGTAACATGAAGCACAATTATATTG
CAAGTATGAACCAGAGATTAATCCAACAGCCCTGTAGAAAAGCCTATCTTACAAATCAACCAG
GAGACACCCATCAGAATGTGGTTACTGAAGCAGGTATAATTCCAATCTAATTATGTTGTT
ATACCAACAATACCCCTGCTTACTGATACTGGTTGCTTTGAAACCTGTTGTTCCAGATGCT
GCATAAAAGTAAAGGAAGAACAAAAACTAGTCAAACACCAGTCTACACTGTGGATTCAAAGAGTA
CCAGAAAAGAAAGTGGCATGGAAGTATAAAACTCATTGACTTGGTCCAGAATTGTAATTCT
GGATCTGTATAAGGAATGGCATCAGAACAAATAGCTTGGAAATGGCTGAAATCACAAAGGATCTG
AAGATGAACTGTAAGCTCCCTTGAGGCAAATATTAAAGTAATTATGCTATTATTCA
TTTAAAGAATATGCTGTGCTAATAATGGAGTGGACATGCTTATTGCTAAAGGATGCACCAA
ACTTCAAACCTCAAGCAAATGAAATGGACAATGCAGATAAGTGTATCAACACGTGGGAGTA
TGTGTGTTAGAAGCAATTCTTTATTCTTACCTTCATAAGTGTATCTAGTCAATGTA
TGTATATTGATTGAAATTACAGTGTGCAAAGTATTACCTTGCTAAAGTGTGATAAAA
ATGAACTGTTCAATATTATTATGGCATCTCATTTCAATACATGCTCTTGTAAAG
AAACTTAACTGTTGTCAGTGAATTCAACACACACAAATATAGTACCATAGAAAAGTTGT
TTCTCGAAATAATTCACTTCAAGTGTGATATAAACCTCTCAAACATTACTAGAGGCAAGGAT
TGTCTAATTCAATTGTGCAAGACATGTGCCTTATAATTATTTAGCTAAACATTAAACAGATT
TTGTAATAATGTAACATTGTTAATAGGTGCATAAACACTAATGCAGTCATTGAAACAAAAGAAG
TGACATACACAATATAATCATATGCTTCACACGTTGCCTATATAATGAGAAGCAGCTCTGA
GGGTTCTGAAATCAATGTGGCCCTCTTGCCTAAACAAAGATGGTTGTCGGGGTTGG
ATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGGTTCCCTAGCTGTATTAGCCTCTG
ACTATATTAGTATAAAAGAGGTATGTGGTTGAGACCAGGTGAATAGTCACATCAGTGTGGAG
ACAAGCACAGCACACAGACATTAGGAAGGAAAGGAACACGAAATCGTGTGAAAATGGTTGG
AACCCATCAGTGATCGCATATTCAATTGATGAGGGTTGCTGAGATAGAAAATGGTGGCTCCTT
CTGCTTATCTCCTAGTTCTCAATGCTTACGCCTTGTCTCAAGAGAAAAGTTGTAAC
CTGGTCTTCATATGTCCTGTGCTCCTTAAACAAATAAGAGTCTTGTGTTCTGGGGAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSLLLGAALLCGHGFRRVSGQKVCFADFKHPCYKMAFHELSSRVSFQEARNACESE
GGVLLSLENEAEQKLIQMLQNLTKPGTGISDGDFWIGLWRNGDGQTSGACPDLYQWSDGSNSQ
YRNWTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSKGRTKTSQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217